Evaluating the Impact of a Universal FRIENDS For Life Programme on Emotional Distress and Academic Self-Perceptions

Julie Paul

Thesis submitted to the University of Nottingham for the degree of Doctor of Applied Educational Psychology, May 2011

ABSTRACT

'FRIENDS For Life' (Barrett, 2004) is a ten-week programme for children aged 7 – 11 years, based on cognitive behavioural principles, designed to teach coping skills and techniques to manage anxiety and depression. This study describes an evaluation of a universal programme, delivered to a class of Year 5 children in a school in a socio-economically disadvantaged community located in the East of England.

A review of literature, combining narrative and systematic approaches, presents what is known about the development of emotional distress and academic self-perceptions in children, underpinned by the principles of Social Cognitive Theory (Bandura, 1986). Evidence for the effectiveness of Cognitive Behavioural Therapy with children is critiqued, with specific attention to the FRIENDS programme delivered as a universal intervention.

A quasi-experimental non-equivalent groups design (intervention group and wait-list control) was employed to evaluate the impact of the programme upon children's levels of emotional distress, their academic self-perceptions and teacher ratings of pupil behaviour. Pre and post-test measures comprised the Paediatric Index of Emotional Distress, (O'Connor et al, 2010), the Myself-As-Learner Scale (Burden, 1998) and the Strengths and Difficulties Questionnaire (Goodman, 1997).

A change score analysis revealed statistically significant reductions in self-reported levels of emotional distress and teacher-rated hyperactivity for the intervention group in comparison to the control group. Both groups showed significantly improved overall behaviour and prosocial skills. There was no evidence of a significant change between or within groups for academic self-perceptions. The limitations associated with quasi-experimental designs are highlighted, together with the difficulties of operationalising abstract constructs such as 'emotional distress' and 'academic self-concept.'

The results are discussed in relation to the theoretical and methodological implications highlighted in previous chapters. Particular attention is paid to the significance of contextual influences operating in concert with the programme components in mediating outcomes. Implications for future research and the role of the Educational Psychologist supporting universal therapeutic programmes in schools are discussed.

ACKNOWLEDGEMENTS

The author would like to thank:

- Anthea Gulliford for her dedicated mentoring and supervision throughout this project.

- Her colleagues at Northamptonshire Educational Psychology Service for their valued support and understanding.

- The staff, children and parents of the participating school for their enthusiasm in helping to make this project such an enjoyable and worthwhile experience.

- Her parents for their unconditional love and for giving up so many days to support our family through this challenging time.

- Her husband for his eternal patience and expert ICT skills.

- Her son, Leo, for forgiving his mummy for not being there.

CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

1 CHAPTER ONE: INTRODUCTION

Introduction to Thesis

This thesis evaluates the impact of the FRIENDS For Life programme (Barrett, 2004) on reducing and preventing emotional distress and enhancing academic self-perceptions in a class of Year 5 children. It also includes teacher ratings of children's behaviour to complement an understanding of the intervention's impact within a universal academic context. The study fulfils a number of purposes, encompassing national, regional and local priorities in policy and educational psychology practice:

i. To contribute to the growing body of knowledge about universal mental health prevention programmes, identified through the previous government's agenda on mental health and psychological well-being (DfES, 2004).

ii. To contribute to the evaluation of FRIENDS, part of the Targeted Mental Health in School's initiative in the researcher's Local Authority.

iii. To supply data towards the University of Nottingham's contribution to the Development and Research (D&R) programme in Educational Psychology. This aims to aggregate outcomes from trainee research conducted nationally, in relation to priority topics under three key themes: developing psychological wellbeing, promoting social inclusion and enhancing educational outcomes. The present study addresses the topic of therapeutic interventions.

iv. To supply data and knowledge about the effectiveness of the programme to the participating school, to support their development of a tailored Personal, Social and Health Education (PSHE) Curriculum.

1.1 Setting the Scene: The National Focus on Well-being

In September 2009, at the inception of this project, educational and social policy in the UK was driven by the Every Child Matters (ECM) agenda (DfES, 2004). Enshrined in the Children Act of 2004, these reforms aimed to improve the outcomes for all children and young people, by providing opportunities to: be healthy, stay safe, enjoy and achieve, make a positive contribution and achieve economic well-being.

At the same time, a survey of the mental health of children and young people in Great Britain provided robust evidence that around 10% aged between five and sixteen had a mental health difficulty that was associated with considerable distress and interference with personal functions, including social aptitude, coping strategies and learning (Green et al, 2004; DCSF 2008). Outcomes indicated that individuals living in families with a low income in areas classed as "hard pressed," were more prone to developing a diagnosable 'disorder', although the correlative nature of these factors was emphasised. A further report (Davidson, 2008) highlighted the lack of consistent data on the prevalence of 'lower level' mental health problems that do not meet the criteria for clinical 'diagnosis,' and pointed to the associations between mental health outcomes and poor educational attainment, absence and exclusion from school and lack of friendship networks. Anxiety in particular has been associated with childhood and adolescent difficulties, (Beesdo, Knappe and Pine, 2009), although there is considerable variation in reported prevalence rates, (6-10%, Carr, 2006; 10-20%, Barrett and Pahl, 2006).

A plethora of political documentation accompanied these findings, (DfES, 2001; DoH, 2004); including the 'Targeted Mental Health in Schools' (TaMHS) initiative which aimed to "transform" the way that mental health services were delivered to primary-aged children through early evidence-based preventive work at the individual, targeted and universal levels, (DCSF, 2010, *p4;* see also Appendix 1a). This is the context in which the researcher's Educational

Psychology Team undertook the FRIENDS For Life programme, which provided the opportunity and inspiration for the present study.

1.2 Rationale

FRIENDS For Life is a ten-week programme, based on cognitive behaviour therapy (CBT) principles, designed to alleviate and prevent anxiety and depression for children aged between 7 – 11 years (Barrett, 2004). Research on FRIENDS delivered as a selective or indicated intervention shows generally consistently positive effects on anxiety, self-esteem and social competence (Bernstein et al, 2005, Liddle and MacMillan, 2010). There remains little evidence, however, about the effectiveness of the programme delivered as a universal intervention in UK schools. Universal preventative interventions target whole populations, as opposed to individuals who have been identified as having a high-risk status (Lowry-Webster, Barrett and Dadds, 2001). This is an important area for investigation, considering the under-identification of children with 'lower level' mental health difficulties (Davidson, 2008) or an absence of externalizing problems (Briesch, Hagermoser Sanetti and Briesch, 2010), and the observation that long waiting lists for specialist mental health services often have adverse effects on children and their families (Kurtz, 2004).

The present study evaluates the impact of FRIENDS, delivered as a whole class intervention, on the reduction and prevention of emotional distress and is the first to specifically measure academic self-perceptions. The rationale for this connection is that children will apply the self-regulatory and coping skills taught through FRIENDS across academic situations, which will in turn impact positively on their sense of themselves as learners (Frydenberg, 2008). The researcher's interest lies in the hypothesised links between children's emotional regulation and their self-beliefs, and the impact of these upon their academic attainments, themes which have received increasing support in educational policy, (DCSF, 2005; OFSTED, 2010). However, as a former teacher, the researcher is equally aware of the tensions that are created for

staff under pressure from competing agendas in schools. As Shoenfeld and Janney (2008) summarise:

"unless anxiety interventions can also provide a corresponding increase in academic achievement, they are unlikely to be considered of any great importance in school settings." (p598)

The study is one of the first evaluations of a universal FRIENDS programme in the UK to include a control group in its design and is therefore well-placed to address some of the methodological limitations of previous studies, including whether reductions in emotional distress can be attributed to the intervention or to the passage of time (Stallard, 2010).

By adopting a quasi-experimental approach, which preserves the intact nature of the classes, the study also provides a distinctive opportunity to explore the ecological context of delivery. Through an analysis of teacher-reported behaviour and detailed observations about programme implementation, the study begins to explore the identified 'gaps' in understanding the processes and mechanisms underlying effective delivery (Briesch et al, 2010). Suggestions are finally made about how these observations might inform a model for the Educational Psychologist's (EP's) role in supporting CBT in schools.

1.3 Summary of Chapters

Chapter One has introduced the topic, themes and rationale for this thesis, including the researcher's personal interest in the links between emotional distress, coping and educational achievement.

Chapter Two reviews what is known about current perspectives on mental health; childhood emotional distress; academic self-perceptions; cognitive behavioural therapy for children and the effectiveness of the FRIENDS programme when applied as a universal intervention. Social Cognitive Theory (SCT, Bandura, 1986) is introduced as a framework within which to explore

the links between these phenomena. The chapter concludes with the research questions and hypotheses arising out of the literature review.

Chapter Three presents the methodology for this study. Current debates around ontology and epistemology within educational research are reviewed as a foundation for explaining the researcher's chosen position and design. The complexities of researching social phenomena in schools are highlighted, together with the practical and ethical considerations associated with evaluative research in educational psychology.

Chapter Four presents the data analysis from the self- and teacher-reports obtained in this study. The complications associated with quasi-experimental designs are discussed, alongside the associated rationale for conducting the chosen analysis. Both descriptive and inferential statistics are presented to address the research questions.

Chapter Five constitutes a discussion of the results in relation to the research questions. The researcher debates the extent to which any observed changes can be attributable to the intervention, and highlights the importance of analysing and monitoring contextual factors to identify possible 'mechanisms' in successful implementation. SCT is revisited to model how the EP might use a universal cognitive behavioural intervention as a vehicle for applying psychology to promote mental health at both the individual and systemic levels. Suggestions for future research are made in light of this study's limitations.

In **Chapter Six**, conclusions are drawn about the evidence for the effectiveness of FRIENDS upon emotional distress, academic self-perceptions and behaviour within the context of this study and how the findings contribute new knowledge in this field of research.

2 CHAPTER TWO: LITERATURE REVIEW

Introduction

This chapter reviews the literature relating to the present research questions and supports a rationale for the researcher's hypotheses. These propose that FRIENDS For Life, a therapeutic cognitive behavioural (CBT) programme, will have beneficial effects on children's self-reported levels of emotional distress, their academic self-perceptions and their teacher's perceptions of their behaviour. Contemporary knowledge about these constructs and the impact of CBT programmes with children are critiqued and set within the national context of mental health in education. Studies investigating the effectiveness of FRIENDS as a universal intervention are evaluated in detail to summarise what is known so far about its strengths and limitations.

The scope of the review covers:

i. Models of mental health and the paradigms within which the Educational Psychologist (EP) may intervene.

ii. The aetiology and maintenance of emotional distress (in particular, anxiety) and its impact upon scholastic functioning.

iii. Two key constructs associated with academic self-perceptions (self-concept and self-efficacy, Burden, 1998a) and how these impact upon academic achievement.

iv. The application and effectiveness of CBT with primary-aged children.

v. The effectiveness of universal applications of FRIENDS in relation to emotional distress and academic self-perceptions.

A narrative review, informed by a non-systematic search, is provided for areas one to four. A full systematic search was carried out to identify individual studies involving universal applications of FRIENDS, which will be detailed in Section 2.5. (Details of specific search strategies are available in Appendices 2a and 2b).

2.1 Intervening With Children's Mental Health

As this project was framed within the mental health agenda, it is important to explore the definitions and paradigms that have influenced the researcher's design and interpretation of outcomes. Some key themes emerging from contemporary research into mental health will be outlined to consider how the present intervention fits within this agenda. The social cognitive perspective (Bandura, 1986) is suggested as a helpful framework within which to explore the effects of the present CBT intervention within an educational setting, as it integrates both internal and contextual factors in the explanation of outcomes. The researcher considers this an important approach for hypothesizing about the processes that might be operating to produce change (Kazdin, 2007). Finally, consideration is given to some contemporary thinking around 'risk' and 'coping' factors that have influenced the national agenda, and how these might affect children's experience of emotional distress and sense of academic competence.

2.1.1. What is mental health?

Weare (2004) highlights the diversity of terms and definitions that have been applied to mental health, including emotional wellbeing, emotional intelligence, psychological wellbeing and mental health problems and disorders. Authors have identified that the term of reference depends on the dominant paradigm within the profession; with education, health, social care and youth justice systems each having their own unique way of framing mental health problems (Weare, 2004; Davidson, 2008). Furthermore, Frederickson, Dunsmuir and Baxter (2009) emphasise how the term mental health is sometimes avoided because of its association with stigmatizing ideas about 'mental illness.' Acknowledging this complexity, the World Health Organization's (2001) explanation of mental health is offered as a helpful working definition that reflects the researcher's assumptions:

"A state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community."
(*p1*).

2.1.2 The concept of mental 'disorder'

When intervening in schools, EPs are faced with the challenge of framing their work within these competing paradigms of mental health. Adopting the medical model assumes an approach subscribing to concepts of diagnostic classification and this has been considered by some, when judiciously applied, as a powerful tool for developing 'treatments' to address particular conditions (Scott, 2002). From this perspective, problems such as anxiety can be categorised, (for example, social phobia, specific disorder, generalised anxiety disorder), and diagnosed against various criteria using the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, American Psychiatric Association, 2000), or the International Classification of Diseases (ICD-10, WHO, 1992). This involves assessing the presence of 'symptoms' over set timescales (Carr, 2006).

Others have challenged these assumptions, however, proposing that an emphasis on 'treating' individual pathology ignores the complexity of problem situations with respect to mental health (Tew, 2005). Tew (2005) advocates "a holistic approach which helps to make links between thoughts, feelings and behaviours, and the realities of people's social and personal experience," (*p25*). Williams (2005) similarly argues that the concept of 'disorder' lacks construct validity and distorts psychological reality. He notes that 'diagnosed' individuals often show idiosyncratic, heterogeneous combinations of problems and responses that cannot be neatly categorised, an observation that is supported by the high comorbidity rates for anxiety 'disorders' (Carr, 2006). Williams (2005) proposes that psychological problems should be viewed as dimensional rather than dichotomous and that successful intervention depends on the appraisal of specific problem responses in context. This is supported by the observation that current definitions of mental health and

'psychopathology' are highly biased towards western, middle class conceptions; yet research suggests that cultural developmental norms and expectations vary considerably according to the culture, family and society to which an individual belongs (Barrett, 2000).

The 'dimensional' versus 'categorical' debate continues to receive much attention and certain authors working within the medical paradigm now suggest that pooled research supports a continuous distribution of emotional difficulties rather than discrete dichotomies (Watson, 2005).

Because this study's purpose was to evaluate the impact of a *universal* FRIENDS programme on *group levels* of emotional distress, the researcher considered that the dimensional approach was a more fitting paradigm within which to consider the intervention. The study does not attempt to address individual symptomatology from a clinical perspective, but considers how cognitive, behavioural and environmental factors interact to affect anxiety levels in a classroom setting. The social cognitive viewpoint that illuminates this position will now be described.

2.1.3 Social Cognitive Theory (SCT)

SCT (Bandura, 1986) emerged as a recurring theme throughout this literature search, linking theories of mental health, emotional distress and academic self-perceptions. The theory will be revisited throughout this thesis, but for the present, a brief introduction to Bandura's conception will be outlined as a framework to consider the constructs under scrutiny. A key feature of SCT which reflects the assumptions of the present CBT intervention, is the human capacity for self-reflection, through which people "make sense of their experiences, explore their cognitions and beliefs, engage in self-evaluation and alter their thinking and behaviour accordingly," (Schunk and Pajares 2009, *p36*). Bandura's model of reciprocal determinism (see ***Figure 2.1***), attempts to explain how personal factors (cognitive, affective, biological), environmental factors and behaviour interact to produce outcomes.

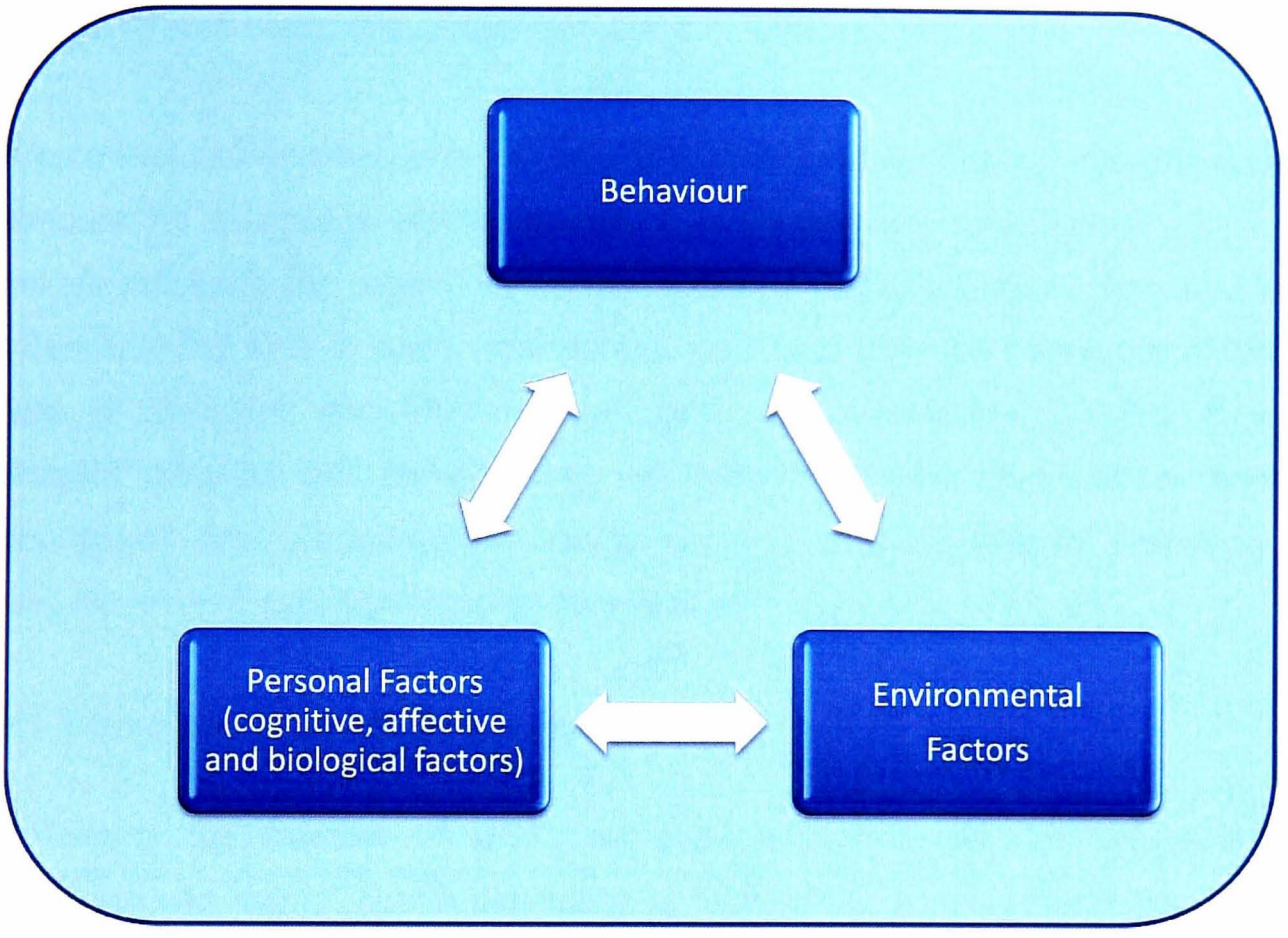


Figure 2.1: Bandura's model of reciprocal determinism.

It is worthy of note that this transactional relationship between environmental and internal factors underpins many contemporary models of problem analysis in educational psychology (Morton and Frith, 1995; BPS, 2002).

2.1.4 Self-Efficacy Theory

Also pertinent to the present analysis is the related theory of self-efficacy (Bandura, 1977), which proposes that the beliefs people hold about their capabilities in relation to designated tasks powerfully influence the ways in which they behave (Usher and Pajares, 2008). Although there is ongoing debate about the nature, definition and measurement of Bandura's constructs, (Pajares, 1996a; Bong and Skaalvik, 2003), authors concur that his theories have made a significant contribution to the understanding of children's functioning in academic contexts, (Bong and Skaalvik, 2003; Schunk and

Pajares, 2009). The following quotation links the themes of self-efficacy, emotional well-being and attainment that are central to this study:

"Perceived self-efficacy refers to beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments... Such beliefs influence the course of action people choose to pursue, how much effort they put forth in given endeavours, how long they will persevere in the face of obstacles and failures, their resilience to adversity, whether their thought patterns are self-hindering or self-aiding, how much stress and depression they experience in coping with taxing environmental demands, and the level of accomplishments they realise." (Bandura, 1997, p3).

2.1.5 Risk and Protective Factors

Reflecting the themes of SCT, evidence suggests that an individual's proneness to mental health difficulties is determined by a complex interplay between intrapersonal, interpersonal and environmental factors (Carr, 2006; Newman, 2004). Identified 'risk factors' include socio-economic deprivation, low self-esteem, few problem-solving skills, fear of failure and social exclusion (Carr, 2000; Dunsmuir, 2010). The outcomes of a study by Fergusson and Lynskey (1996) suggest that risk factors operate in an interactive, cumulative fashion and this has been supported by further longitudinal studies. Schoon (2006), for example, notes that cumulative adversity produces a negative chain effect on levels of academic adjustment, which in turn impacts on subsequent attainments. However, although there is a strong relationship between exposure to hardship and developmental outcomes, there is substantial diversity in the way that individuals respond (*ibid, p74*).

Frederickson et al (2009) concur that exposure to risk does not inevitably produce negative outcomes and protective factors which support growth and development can act to "buffer the effects of adverse experiences," (*p 1*). Carr (2000) identified such factors as high self-esteem, good problem-solving skills, a love of learning, good communication skills, having close friends and the delivery of a comprehensive Personal Social and Health curriculum as

being effective in this regard. Newman (2004, *p4*) suggests that a child's 'resilience', that is their capacity to "resist adversity, cope with uncertainty and recover more successfully from traumatic events or episodes," arises from multiple, dynamic interactions between these factors and others such as the child's temperament. This observation has important implications as it suggests that interventions to promote resilient adaptation may help to shape the course of development, although this will be affected by the individual's capacity to engage and respond effectively (Newman, 2004).

2.1.6 Coping

According to Frydenberg (2008), coping can be broadly defined as the thoughts, feelings and actions used by an individual to deal with problematic situations. In line with SCT, Frydenberg (2008) represents coping as multi-dimensional, dynamic and context-dependent. Evidence suggests that children spontaneously apply coping strategies such as problem-solving, support-seeking, rumination, escape and distraction, and that these interact with their academic and social functioning, adjustment to stressful events, internalizing and externalizing behaviour, well-being, competence and resilience, (Zimmer-Gembeck and Skinner, 2011). In an integrative review of studies, Zimmer-Gembeck and Skinner (2011) found that patterns of coping became more differentiated with age, with the development of language and metacognitive capacity during middle childhood increasing the sophistication of strategy selection. This evidence supports proponents of CBT who advocate its use with children and young people to improve coping capacity (Newman, 2004). However, Zimmer-Gembeck and Skinner (2011) emphasise the limitations regarding design and analysis of developmental change in current coping research, while critics highlight the complexities of applying adult models of CBT to children (Stallard 2002; see Section 2.4).

2.1.7 Summary

In this section some definitions of mental health have been explored. It was proposed that when intervening with mental health in schools, dimensional models of mental health that acknowledge the interaction between personal, environmental and behavioural factors offer EPs a more dynamic framework for exploring developmental outcomes than static, categorical models. SCT, (Bandura, 1986), has been introduced as a framework supported by research linking the mental health and academic domains. It was suggested that process-focused interventions such as CBT may be effective in strengthening children's coping strategies, which enhance resilience to adversity, although there are limitations in our understanding of the interaction between coping and developmental change.

2.2 Emotional Distress in Childhood

Introduction

The present study examines the impact of a universal CBT intervention on the construct of emotional distress, which is defined here as comprising anxiety and depression (O'Connor et al, 2010). As depression is considered to be less common in pre-adolescents (Carr 2006), and the scope of this review does not permit a detailed exploration of both constructs, this section will focus predominantly on the literature related to anxiety, with brief reference to its links with depression at the end. Whilst acknowledging the breadth of theoretical approaches in this field, (Carr, 2006), the researcher has chosen to focus specifically on cognitive and transactional theories because these bear direct relevance to the current CBT intervention as applied in a universal context. Factors associated with the development, maintenance and impact of anxiety will be highlighted in order to contextualise how the present intervention might address them. Reflecting the argument in section 2.1, anxiety will be considered as a continuum and reference to 'disorders' will be generally avoided.

2.2.1 Phenomenology of Anxiety

In an influential model, Lang (1979, cited in Ollendick, Shortt and Sander, 2005) represented anxiety as a tripartite phenomenon involving *cognitive appraisal, physiological arousal and avoidant behaviour.* Bandura (1977) challenged this multifaceted definition as conceptually problematic and argued for the disaggregation of the components in order to theoretically test the relationships between them. Anxiety is thus conceptualized as "an emotion of fright indexed by physiological arousal or subjective feelings of agitation" (Bandura, 1997, *p138*). Contemporary definitions continue to explain the construct in terms of physiology and emotion, (OED, 2002; Amstadter, 2008), but the components of the tripartite model continue to form the foundation of many current CBT programmes (Appendix 2c).

2.2.2 Origins of anxiety

Allen and Rapee (2005) emphasise that formulating appropriate interventions depends on understanding the possible factors implicated in the origin and maintenance of anxiety.

Cognitive Explanations

Cognitive theorists, (Beck, Emery and Greenberg, 1985) propose that schemas are formed in response to threatening or stressful experiences. According to this view, childhood anxiety, akin to adult models, is associated with 'distorted cognition', with variables including negative thinking, worrying, causal attributions and biased attention and memory processes, (Prins, 2001, *p23*). A range of evidence suggests that deficient or distorted cognitive processing is associated with psychological problems in children and that these impact upon affect and behaviour, although it has been questioned whether theoretical models based on the dysfunctional cognitions of adults can be extrapolated to children (Stallard, 2002). Prins (2001) also cautions that a causal relationship between distorted cognition and maladaptive behaviour is far from established and the possibility remains that some cognitive correlates result *from* the anxiety, perhaps indicating a more circular relationship.

Social Cognitive Explanations

SCT proposes that perceived efficacy to exercise control over potentially threatening events plays a key role in anxiety arousal and avoidant responses (Bandura, 1997). According to this view, enhancing perceived control of a threatening situation (rather than using techniques such as relaxation) is viewed as the most effective way to obviate emotional distress (Bandura, 1997). Bandura's theory of emotional regulation has had a significant impact upon educational research and practice (Usher and Pajares, 2008), and will be critiqued further in section 2.3.

Transactional Explanations

Current perspectives propose a dynamic, transactional model of anxiety development that is highly consistent with the tenets of SCT (Ollendick et al, 2005). According to this model, multiple developmental pathways emanate from the relationship between genetic, constitutional, physiological, behavioural, psychological, environmental and sociological factors (*ibid, p355*), a theory that has received increasing support in research. Lau, Eley and Stevenson (2006), for example, demonstrated that the relationship between state and trait anxiety, (where trait anxiety is expressed through levels of state anxiety under threatening circumstances), represents a process of interplay between genetic vulnerability factors and environmental stressors.

Other studies have highlighted the reciprocal role of social factors in the maintenance of anxiety. Allen and Rapee (2005), for example, discovered that children with "inhibited temperaments" selectively chose friends modelling anxious behaviours, while Biggs, Nelson and Sampilo (2010), found that emotional distress in adolescents was significantly mediated by peer relations. However, contemporary studies continue to stress the importance of identifying the factors and mechanisms implicated in emotional distress in temperamentally at-risk populations (Degnan, Almas and Fox, 2010).

The difficulty of testing specific components of transactional models with whole group variable-centred approaches has been raised by Zahn-Waxler, Klimes-Dougan and Slattery, (2000). While the researcher acknowledges this view in relation to investigating 'clinical' subtypes of anxiety and depression, Chapter Five addresses whether whole group designs are able to illuminate understanding of transactional factors implicated in reducing levels of emotional distress in classroom settings.

2.2.3 Developmental Trajectories

According to Amstadter (2008), anxiety can be an adaptive emotional response by triggering anticipatory problem-solving thoughts, but it can also consume attentional resources and lead to feelings of helplessness and withdrawal, and under these circumstances may be considered maladaptive (Ollendick et al, 2005).

It has been noted that 'normal' anxieties follow a typical developmental sequence through childhood into adolescence (Carr, 2006). Of particular relevance here is the observation that during middle childhood, as children begin to make social comparisons, the focus shifts from early worries about fantasy creatures, separation and the dark, for example, to those associated with academic and athletic performance. This has prompted the argument that the treatment of anxiety in school settings is particularly relevant as many childhood anxieties arise in response to school-based stimuli (Briesch et al, 2010).

Studies suggest that *maladaptive* anxiety interferes with many aspects of functioning including interpersonal relationships, social competence and school adjustment, (Barrett, Lock and Farrell, 2005; Stallard, 2010). Further research has highlighted that, when left untreated, anxiety problems in childhood lead to an increased risk of developing further mental health problems in young adulthood, (Pine et al, 1998); impaired social adjustment (Last, Hansen and Franco, 1997) and an increased propensity to depression, (Frydenberg, 2008). As this study is concerned with the impact of emotional distress on academic functioning, the next section will examine evidence for this relationship.

2.2.4 The Impact of Anxiety on Scholastic Functioning

A number of cross-sectional and longitudinal studies have supported the association between anxiety and impaired scholastic functioning, (Langley et al, 2004; Massetti et al, 2008). Research suggests that anxiety interferes with

cognitive performance because elevated physiological arousal leads to a narrowing of attention focus, impaired concentration and disturbance in memory recall, (Wood, 2006), while recent studies have suggested a reciprocal relationship between students' environments and positive emotions (Frydenberg, 2008). In a systematic review of anxiety treatment studies involving children with emotional and behavioural disorders, Shoenfeld and Janney (2008) found that seven out of eight cases reported a significant negative effect between anxiety and performance. Further research has highlighted the longitudinal adverse effects of anxiety on educational prospects, (Duchesne et al, 2008; Van Ameringen, Mancini and Farvolden, 2003), although this latter study was limited by its retrospective design and clinic-based sample.

Despite these associations, studies into the specific mechanisms underlying the relationship between anxiety, coping and performance have produced ambiguous results (Prins 2001). Variations have also been found in the range and types of coping strategies applied, according to age, gender and whether students are reported to have emotional and behavioural problems, with the latter group displaying less problem-focused and emotion-focused coping, and increased stress and maladaptive coping (Hampel and Petermann, 2005). These findings suggest that individual factors play a prominent role in the relationship between self-regulation, coping and performance, a theme that is continued below and revisited again in section 2.3.

2.2.5 Comorbidity With Externalizing Problems

Comorbidity refers to the existence of two or more discrete problems perceived in the same individual at the same time (Baldwin and Dadds, 2008). There is substantial evidence to support the comorbidity of anxiety with externalizing problems such as behaviour and attention difficulties (Carr, 2006; Costello et al, 2003). Furthermore, it has been noted that school functioning in children with high anxiety is negatively impacted by the presence of attention problems (Hammerness et al, 2010). However, the

relationship between these constructs is complex and the presence of one does not necessarily predict the development of the other, (Baldwin and Dadds, 2008; Hammerness et al, 2010). Possible explanations include that the presence of anxiety may moderate the expression of externalizing problems or that a third variable such as 'negative affectivity,' (defined by Clark and Watson, 1991, as the degree to which a person "encompasses various aversive states including upset, angry, afraid, sad, scornful, disgusted and worried", *p 321*), is implicated in aspects of both (Baldwin and Dadds, 2008). Meanwhile, although evidence supports the use of behavioural treatments for externalizing problems, the benefits of adding a cognitive component are still disputed (Wolpert et al, 2006).

While a detailed exploration of these issues is beyond the scope of the current review, the present study affords the opportunity to evaluate teacher-rated perceptions of pupil behaviour alongside self-reported emotional distress. The researcher considered this a useful inclusion to explore the processes that might be occurring as a result of the intervention (or otherwise) within the classroom environment.

2.2.6 Depression

The second component implicated in the construct of 'emotional distress' is depression (O'Connor et al, 2010). As the literature on pre-adolescent depression is less extensive and is closely associated with the anxiety findings, a brief summary of relevant points will be included here.

According to Carr (2006), depression in youngsters under 18 ranges from 2% to 9% and is more common in adolescents than pre-adolescents.
Depression is characterized by persistent sadness, a loss of interest or pleasure in activities, lethargy and social withdrawal, and can be associated with attention problems, feelings of failure, hyperactivity and aggression (Frydenberg, 2008; Barrett, 2004). Evidence increasingly supports a temporal, if not causal, relationship between anxiety and depression in children and young people (Barrett, 2004).

Cognitive theories of depression highlight the activation of negative schemas associated with loss, (Beck, 1979, cited in Rait, Monsen and Squires, 2010). According to this view, the interpretation of situations is influenced by depressive cognitive distortions, which affect automatic thoughts (Carr 2006). Evidence suggests that avoidant or aggressive coping styles are linked closely with the development of depression, (Seiffge-Krenke, 2000), while productive coping is associated with lower prevalence, (Kraaij et al. 2003). As with anxiety, depression has been associated with difficulties in concentration, social relationships and school performance (Fröjd et al. 2008). By implication, CBT techniques that aim to challenge negative assumptions have received attention and some support in literature (see Section 2.4). It is important to note, however, that much of the research has focused on adolescents and studies into the effects of CBT on younger depressed children remains limited.

2.2.7 Summary

In this section a developmental approach to the study of emotional distress has been advocated which highlights the complex interplay between individual and environmental factors. The role of social learning and cognition as a mediator in developmental pathways has been highlighted. There is considerable evidence to support a relationship between emotional distress and impaired academic functioning although the specific mechanisms underlying the relationship between anxiety, increased coping and performance are a subject of ongoing research. Studies on depression reflect similar associations with non-productive coping styles and poor school performance.

2.3 The Self As A Learner

Introduction

This study poses the question of whether a universal CBT programme is successful in reducing emotional distress and enhancing the academic self-perceptions of Key Stage 2 children. Domains relevant to the area of emotional distress were examined in the previous section. The focus here will be upon studies of academic self-perceptions in order to provide a rationale for how the components of the programme might impact upon these. Following Burden (1998a), the researcher draws specifically on theories relating to academic self-concept and self-efficacy, which form the basis of the measure used in this study. The section focuses particularly on the complexities of defining and operationalising academic self-perception constructs and also considers evidence for their malleability through intervention. The relationship between self-perceptions and academic achievement will also be outlined to emphasise the potential impact of intervening within this domain.

2.3.1 The Influence of Self-Perceptions

Burden (1999) describes how the search for ways of elucidating people's visions about themselves has a history "almost as long as psychology itself" (*p1*). He attributes this to an ongoing fascination with the "commonsense notion" that how people feel about themselves must be linked in some way with their perceived successes and failures. In an extensive study to explore this hypothesis, Oosterwegel and Oppenheimer (1993) conclude that people not only process information about themselves, but use the resulting cognitive representations to guide their future actions:

> *"Self knowledge is not organized in a static structure but in an active, dynamic structure that continuously interprets and organizes self-relevant actions and experiences and consequently, mediates and regulates behaviour and affect." (pxi)*

This view of the self as a multidimensional, dynamic phenomenon, involving both core constructs and socially constructed elements, echoes the assumptions of SCT and implies that self-perceptions are malleable through intervention (Smiley and Dweck, 1994). Reflecting the national context described in section 1.1, research over the past couple of decades has focused increasingly on the impact of enhancing self-perceptions on mental health and school performance.

2.3.2 The Dilemmas of Studying Self-Perceptions

Authors have drawn attention to the abundance of terms that have been applied when defining aspects of the self (Shavelson, Hubner and Stanton, 1976; Pajares 1996b). According to Rayner (2001), terms such as self-esteem, self-image, self-confidence, self-evaluation and self-efficacy have frequently been used interchangeably, leading to the absence of a universally accepted definition of 'self-concept', for example. This phenomenon may be attributed to the fact that studies of self-perceptions have been shaped by the convergence of a number of (sometimes conflicting) psychological paradigms, including developmental, phenomenological, experimental and social constructionist perspectives (Rayner, 2001). These variations present issues for the researcher because they raise epistemological questions about how constructs of the self can be meaningfully appraised.

The difficulties associated with deciphering, defining and measuring these highly analogous hypothetical constructs were illustrated by Hansford and Hattie (1982). In a meta-analysis of 128 studies they found only moderate associations and a large range in the relationship between self-measures, performance and achievement. Implicated in this variation were factors such as the self-term used, the type and name of the self-test and the reliability of

both the self-ratings and performance/achievement measures; discrepancies that have also been noted in subsequent reviews, (Craven, Marsh and Burnett, 2004). Although Burden (1999) has argued that increased sophistication in test development has improved the robustness of self-perception scales, this evidence highlights the need for a critical approach when analysing reports of self-perceptions.

2.3.3 The Differentiated Notion of Self

Stemming from early theoretical explanations is the notion that the self comprises different components. James (1890, cited in Rayner, 2001) was one of the first psychologists of the modern era to divide studies of the self into its "constituent" parts, the feelings and emotions they arouse and the actions they prompt (*original, p292*). James differentiated between the "I", the thinking self, and the "ME", the sense of self related to experience, as well as introducing the notion of a hierarchy in the self-concept.

Authors continue to report the influence of James's conceptualizations on contemporary theoretical frameworks (Rayner, 2001; Hacker, Dunlosky and Graesser, 2009). Although some experimental research has attempted to explain self-concept in 'global' or 'unitary' terms, little evidence has been found to validate this theory, (eg. Piers and Harris, 1969). Subsequent researchers have focused increasingly on self-concept as a multifaceted, dynamic system with differentiated, context-specific subcategories (Shavelson, Hubner and Stanton, 1976). It is from this framework that the notion of an 'academic self-concept' has emerged.

2.3.4 A Hierarchical Model of Self-Concept

An influential model of self-concept as a multidimensional, hierarchical structure was proposed by Shavelson et al (1976) and this continues to form the basis of many contemporary investigations (Craven and Marsh, 2008; Brunner et al, 2010; see ***Figure 2.2***). Shavelson and Bolus (1982) employed a structural equation model to produce a theoretical structure with a stable, general self-concept at the apex of a pyramid, supported by academic and non-academic self-concepts underneath. Academic self-concept is then further subdivided into specific subject areas, e.g. Maths, English, and non-academic self-concept is divided into social, emotional and physical subcategories. According to the model, self-concept becomes less stable with increased situation specificity.

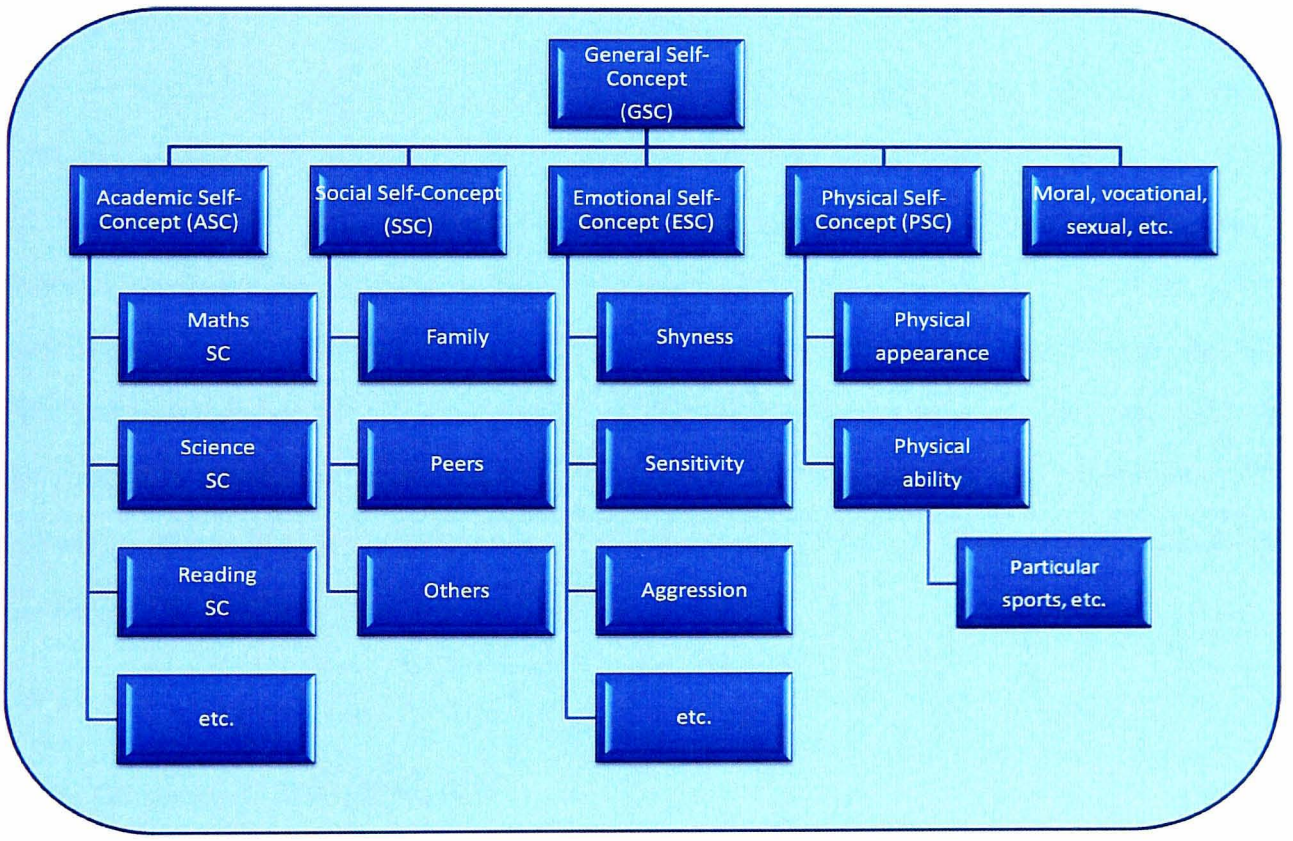


Figure 2.2: A representation of a multi-faceted, hierarchical model of self-concept (adapted from Burden, 1999).

A robust review of evidence supports the construct validity of a multidimensional perspective of self-concept, (Marsh, 2008, *p454*). Marsh (2008) suggests that, for researchers working in academic settings, choosing selected domains of self-concept is associated with greater reliability and predictive value than global measures of 'self-esteem.' The question for the present study is whether the hypothetical domain of 'academic self-concept' is sufficiently differentiated to detect subtle changes in self-perceptions as a result of the intervention. The theory predicts that 'academic self concept' may be only partially receptive to change, being in a central and potentially stable part of the hierarchy (Marsh, 1990).

2.3.5 Academic Self-Concept

Academic self-concept has been defined as a description of one's perceived self, accompanied by an evaluative judgement of self-worth or the knowledge and perceptions that one holds in achievement situations, (Pajares and Schunk, 2001; Bong and Skaalvik, 2003). It has been argued, however, that its usefulness as a construct is dependent upon its ability to explain and predict outcomes (Bong and Skaalvik, 2003). As noted above, the operationalisation of self-concept as a global construct has historically led to methodological weaknesses and poor predictive outcomes. The implication for future research is that any understanding of the impact of children's self-concept upon school functioning must take the effect of domain into account (Craven, Marsh and Burnett, 2004; Marsh, 2008).

2.3.6 Academic Self-Efficacy

The construct of academic self-efficacy is also implicated in the notion of academic self-perceptions (Burden, 1998). According to Bong and Skaalvik (2003), academic self-efficacy refers to an individual's convictions that they can perform given academic tasks at designated levels. The present study considers whether a group CBT intervention is successful in enhancing

academic self-efficacy. This involves examining the effects of teaching strategies to limit self-hindering thoughts, increasing cognitive coping skills, developing goal setting strategies and promoting the development of self-regulatory skills to improve self-efficacy in problematic academic situations.

Sources of Efficacy Beliefs

Bandura (1997) proposed that self-efficacy beliefs are created and developed from the individual's interpretation of four sources of information: mastery experience of previous achievements, vicarious experience, the verbal and social persuasions of others and feedback from emotional and physiological states. According to a recent comprehensive review by Usher and Pajares (2008), mastery experience is the most consistent and influential source of information that students use to inform their self-efficacy beliefs. Correlations between vicarious experience and self-efficacy have proved inconsistent while social persuasions have not proven predictive of self-efficacy across all contexts when other sources are controlled (*p26*). Meanwhile, although SCT predicts a strongly negative relationship between anxiety and self-efficacy, correlations exploring emotional and physiological indices produced a wide range from -0.08 to -0.57. While the authors highlighted methodological limitations in this literature, these results raise intriguing questions about the relationship between arousal and self-efficacy, with the possibility that arousal perceived as a challenge may actually enhance performance (Bandura 1997). The relationship between anxiety and self-efficacy will be discussed further below.

2.3.7 The Relationship Between Anxiety and Self-Efficacy

It was suggested in the previous section that the relationship between physiological arousal and performance is complex. This section considers further the evidence to support a link between enhanced emotional self-regulation and self-efficacy in academic contexts.

Bandura (1988) suggested that people's belief in their capabilities affects how much stress they feel in threatening situations. He argued that perceived efficacy to exercise control over stressors plays a key role in the level of anxiety arousal. Bandura (1988) demonstrated that people who believe they can exercise control over potential threats do not experience such apprehensive cognitions, while those who doubt their ability to cope experience high levels of anxiety and physiological arousal. Bandura's conclusion that strengthening coping efficacy through guided mastery experiences results in lower anxiety symptoms and less avoidant behaviour will be revisited in Chapter Five.

Bandura's theories continue to find support in contemporary studies reporting associations between low academic self-efficacy and anxiety and depression (Muris, 2002); self-efficacy as a mediator between test anxiety and academic achievement (Shunsen and Guiqing, 2009) and the influence of early emotional and behavioural self-regulatory skills on later academic self-efficacy beliefs and literacy achievement (Liew et al, 2008). However, the last authors highlighted the need to differentiate the roles of specific self-regulatory components and to include a sample with more varying academic ability across a wider range of subjects.

2.3.8 The Impact of Self-Perceptions on Academic Achievement

One of the intentions of this study is to contribute to an understanding of whether a CBT intervention may ultimately have a beneficial effect on academic outcomes. It would therefore be important to demonstrate a link between enhanced academic self-perceptions and achievement.

Although debate has arisen over the direction of influence regarding self-concept and achievement, (Craven, Marsh and Burnett, 2004), robust evidence now suggests that domain-specific self-concept predicts subsequent attainment (Marsh and Yeung, 1997; Choi, 2005), while a recent review provides support for a reciprocal effects model (Marsh and Martin, 2011).

Similar strong evidence links the appraisal of capabilities (self-efficacy) with a number of cognitive and motivational processes, and subsequent achievements. Zimmerman, Bandura and Martinez-Pons (1992) found that students' beliefs in their efficacy for self-regulated learning affected their perceived self-efficacy for academic achievement, which in turn influenced their goals and subsequent attainments. In a seminal meta-analysis of studies, Multon, Brown and Lent (1991) found positive and statistically significant relationships between self-efficacy beliefs and academic performance and persistence outcomes. Contemporary studies continue to reveal evidence for the relationship between perceived self-efficacy and academic continuance and achievement, (Caprara et al, 2011), concluding that SCT "provides guidelines for enhancing students' efficacy to regulate their learning activities" (*ibid*, *p78*).

Others have highlighted complexity in investigating the psychological mechanisms underlying the relationships between self-efficacy and achievement, however. Schunk and Pajares (2009) suggest that the issue of calibration (ie. how well self-efficacy relates to actual performance) is complicated by a number of instructional, social and cultural factors in schools. Furthermore, as with studies of self-concept, the predictive power of the self-efficacy construct appears to increase with appropriate levels of specificity (Pajares 1996a). The implication is that the assessment of general self-efficacy would reveal little about an individual's functioning in a particular subject domain (Usher and Pajares 2008). Meanwhile, Usher and Pajares (2008) note methodological and conceptual anomalies in some of their selected studies; for example, the strength and influence of the sources of self-efficacy varied according to contextual factors such as gender, ethnicity

and academic ability. They also questioned the use of objective measures in self-efficacy research that may mask the meanings individuals attach to their performance. In conclusion, although methodological limitations should be acknowledged, there is sufficient growing evidence to suggest a correlational, if not causal link between self-perceptions and school achievement, with a reciprocal effects model gaining increasing support.

2.3.9 Summary

In this section, evidence has been presented for the influence of academic self-perceptions on school achievement. This implies that enhancing children's views of themselves in relation to academic functioning could be an important focus for intervention. Evidence suggests that hierarchical conceptualisations with attention to specific domains provide the most robust models for educational research, rather than more generalised, global measures of self-esteem. The researcher hypothesises that an intervention aimed at enhancing children's coping skills and self-efficacy in school will have a beneficial effect on their academic self-perceptions. The theoretical and methodological complexities of researching and measuring hypothetical self-constructs have been discussed and these will be considered later when the present results are interpreted.

2.4 Cognitive Behaviour Therapy For Children

Introduction

The FRIENDS programme, created to assist children with learning important skills and techniques to manage and prevent anxiety and depression, is based on CBT principles. This section outlines the theoretical underpinnings of CBT and explores some of the debates surrounding its use with children and young people. The section concludes with an evaluation of the evidence for CBT's effectiveness in addressing emotional distress with younger age groups to support the current hypotheses.

2.4.1 The Components of CBT

CBTs constitute eclectic mixes of techniques combining strategies from cognitive and behavioural psychology (Rait, Monsen and Squires, 2010). Graham (2005) differentiates between contemporary narrow definitions of CBT, focusing on specific therapies targeted at modifying thoughts, feelings and behaviours and a more broad definition that encompasses a range of approaches including social skills training, solution-focused and anger management techniques. Rait et al (2010) suggest that the EP must make a judgement about where on this 'continuum' a particular CBT programme may be operating.

The empirical base for current CBT approaches with children is drawn principally from two 'schools': Ellis's Rational Emotive Behaviour Therapy (REBT, Ellis, 1957) and Beck's Cognitive Therapy (CT, Beck, 1979; both cited in Rait et al, 2010). According to Ellis (1980), emotional disturbances are largely created through people's fundamental, inflexible beliefs about activating events and thus the aim of therapy is to challenge the resulting irrational thoughts, feelings and behaviours. Beck's CT approach alternatively posits that psychological difficulties arise from distorted information processing and the maladaptive appraisal of stimuli as described section 2.2.2 (Bolton, 2005), and therapy involves encouraging individuals to investigate the

nature and origin of their negative thoughts. Ellis (1980) argued that REBT differed from other CBT models in its strong emphasis upon philosophical and humanistic elements, although it is now acknowledged that REBT and CBT have had a mutually complementary influence on one another (Dryden and David, 2008).

The FRIENDS programme draws strongly on elements of Ellis's and Beck's theories, as highlighted in Stallard's (2010) summary of its core elements. These include: psycho-education to teach relationship between thoughts, feelings and behaviour, emotional recognition and management training; recognising cognitions (including distortions) in anxiety-provoking situations; challenging negative self-talk with positive coping and anxiety-reducing self-talk; practising new emotional and cognitive skills and self-reward for positive attempts at overcoming worries. Stallard (2010) importantly highlights the differences between teaching these broad skills at a classroom level and the more specialised application of CBT at an individual therapeutic level, although the appropriate level of specialism required to deliver universal CBT is still the subject of some debate (Briesch et al, 2010; Squires, 2010).

While the parsimonious and heuristic value of CBTs has been highlighted, together with their precision and testability, limitations have been noted in the theories' failure to address the developmental aspect of "irrational beliefs" or "faulty information processing" (Sapp, 2004; Rait et al, 2010). This has prompted discussion about whether techniques requiring advanced cognitive and verbal ability are appropriate for use with children (Grave and Blissett, 2004). This issue and the related empirical evidence for the effectiveness of CBT with younger age groups is addressed in the following sections.

2.4.2 Theoretical Considerations Underlying the Use of CBT with Children and Young People

While CBT has consistently been found to be effective in ameliorating adult mental health difficulties, (Butler et al. 2006), its application to children and young people has been more contentious. Much discussion has centred upon whether adult models of CBT can be extrapolated to children; and in particular whether the age and developmental level of the child interacts with the success of the treatment, (Stallard, 2002; Kendall and Choudhury, 2003; Bolton, 2005). Stallard (2002) also raises the importance of considering systemic influences upon the development and maintenance of dysfunctional cognitive processing. Addressing these issues is critical, considering the poor prognosis of childhood mental health difficulties and the current emphasis upon early intervention, as well as the increased interest in universal prevention programmes (see discussion later in this section).

It has been suggested that a developmental perspective offers some insights into whether CBT components might plausibly interact with the physical, emotional and cognitive capabilities of the child (Barrett, 2000). Comprehensive reviews of the developmental evidence relating the use of CBT to children can be found in Grave and Blissett (2004) and Graham (2005) and a brief summary will be provided here. According to Piaget's stage model, preoperational children (2-7 years), whose thinking is dominated by perception, would be incapable of executing the abstract and hypothetical cognition involved in CBT; this would not be fully realised until the formal operational stage in adolescence. Support for this reasoning has been found by Durlak, Fuhrman and Lampton (1991), who discovered that the effect size for children aged between 11 and 13 was almost twice that for those in the concrete operational and preoperational stages, leading them to conclude that cognitive developmental level mediates the outcomes of CBT.

Subsequent research has questioned this analysis and interpretation, however, and it is now widely accepted that the picture of younger children's reasoning abilities is more mixed, (Stallard, 2002; Grave and Blissett, 2004).

There is some evidence that by middle childhood (and in some cases even earlier) children are capable of deploying the kinds of metacognitive strategies required in CBT. These include: theory building; more complex appraisal of thoughts and emotions; understanding that mental states are controllable; and appreciating the concept of 'theory of mind,' (Bolton, 2005; Zimmer-Gembeck and Skinner, 2011). Stallard (2009) describes how children under the age of 7 can distinguish between thoughts, feelings and actions and can appreciate that two people may have different thoughts about the same event, although Flavell, Flavell and Green (2001) found that five year olds were less proficient at making links between thoughts and emotions than eight year olds. Meanwhile, researchers have demonstrated that young children's performance on hypothetical thinking and analogical reasoning tasks can be significantly enhanced by modifying or expanding the question (Grave and Blissett, 2004). Thus while it may be apparent that cognitive reasoning becomes more sophisticated with age, these results suggest that younger children's cognitive competence may be severely underestimated because their performance depends on the methodology and language used, (Shaffer, 1996). Grave and Blissett (2004) conclude: "the implication... is that given clear, simple instructions in the use of these skills, based upon familiar materials from their everyday lives, children may be capable of, and benefit clinically from, cognitive procedures at an earlier age than experimental psychology might suggest," (p406).

2.4.3 Contemporary Evidence for the Effectiveness of CBT on Emotional Distress with Primary-Aged Children

A number of recent critical reviews provide support for the effectiveness of CBT approaches in relieving and preventing emotional distress in children.

Short-Term Effects on Anxiety and Depression

Systematic reviews and meta analyses focusing on studies judged to have generally high standards of methodological rigour have revealed medium to large effects for the short-term impact of CBT on anxiety in children (Compton

et al, 2004; Cartwright-Hatton et al, 2004; James, Soler and Weatherall, 2005).

A review by Merry et al (2004) also provides some support for the effectiveness of psychological interventions in reducing depressive symptoms when compared to non-intervention controls. However, CBT formed only part of this analysis and significant effects were found only in the targeted interventions.

Long-Term Effects

Two studies have indicated that treatment gains from CBT appear durable (James, Soler and Weatherall 2005; Nevo and Manassis, 2009), although the latter authors highlighted the problem of accounting for confounding variables such as maturation.

2.4.4 Universal or Targeted Approaches?

As Stallard (2010) points out, although the results of small-scale, clinic based studies have shown promise in this field, the results cannot necessarily be applied to prevention programmes in schools. Selective reviews have drawn attention to the variable evidence for the impact of universal school-based approaches on preventing depression, for example, (Merry et al, 2004; Spence and Shortt, 2007), although the latter authors reported sufficient ambiguity to merit further research. Subsequent reviews have suggested more positive trends, (Adi et al. 2007; Neil and Christensen, 2009). Results of the latter study, involving a large CBT component including FRIENDS, indicated that 69% of universal trials reported significant differences between intervention and control conditions at post-test, (*ES* = 0.31-1.37).

Another question has been raised about how beneficial universal programmes are for children not at risk of developing anxiety problems. Gallegos (2008) suggests that the underlying philosophy of universal prevention programmes is not just to reduce the chances of future problems occurring but also to

promote competencies that may benefit individuals, regardless of risk status. The present study analyses preventive effects by examining the number of children moving in and out of the 'range of concern' on the emotional distress and academic self-perception measures.

2.4.5 Critique of CBT Literature

All of the above authors have drawn attention to methodological weaknesses and limitations in the studies conducted so far. Cartwright-Hatton et al (2004) identified that the quality of some trial reporting was variable in terms of randomisation, intent-to-treat analysis, statistical power and attrition rates, with all studies scoring poorly in at least one domain. There was also considerable heterogeneity in the types of therapy and the reported effect sizes, with reduced effects for larger studies. Several researchers have commented that more scrutiny is needed around the underlying mechanisms in CBT, including the specific components that achieve therapeutic effects (Compton et al, 2004; King, Heyne and Ollendick, 2005), and others have questioned whether a direct focus on cognitions is essential for reducing anxiety (Stallard 2010). Meanwhile, the need for active comparison groups and further long-term follow-up has been raised throughout this literature (Cartwright-Hatton et al, 2004; Merry et al, 2005; Neil and Christensen, 2005).

While evidence suggests tentative support for the efficacy of CBT approaches with children over the age of six, (Cartwright-Hatton et al, 2004), important areas are highlighted for future investigation. O'Connor and Creswell (2005) point out that the empirical evidence for a developmental model underlying CBT with children remains unverified at present and treatment outcomes appear more related to context, with age being a poor predictor of response. Grave and Blissett (2004) counter that age and developmental level probably do play a mediating role but that it has been difficult to establish relationships, partly due to the lack of sophistication in the assessment and measurement of cognitive function and change in children. Most authors agree that more evidence is needed regarding how the specific elements of CBT interact with

developmental pathways to predict positive or negative outcomes. Furthermore, most of the studies so far have focused on the efficacy of CBT in remediating short-term anxiety symptoms and research into the potential of the approach as a preventive process is still in its infancy (Stallard, 2010). It has also been observed that few CBT studies have actually attempted to evaluate cognitive changes or more central aspects of the child's sense of competence (James et al., 2005 Grave and Blissett ,2004).

2.4.6 EPs Supporting the Delivery of CBT

In their review of the functions and contribution of EPs, Farrell et al (2006) indicate that there is an opportunity to expand the scope of the profession's engagement in therapeutic work, including CBT. Rait et al (2010) point out that EPs have a unique working knowledge of school systems and how these impact on children's learning and behaviour, and suggest that this places them in an "ideal position" to support staff with the delivery of CBT programmes. They also propose that the implementation of more universal CBT interventions reduces the stigma associated with targeted and selected programmes and supports the notion of school as a "therapeutic environment," (*p114*). Not only could this promote the prevention of future psychological difficulties, (Stallard, 2010), but it may also address the concerns associated with long waiting lists for specialist mental health services (Kurtz, 2004). These suggestions will be revisited in Chapter Five.

2.4.7 Summary

At present there are limitations in the evidence for a developmental model underlying the application of CBT to children. However, experimental findings indicate that children at least over the age of six possess the cognitive capability to access CBT-based programmes that are tailored to their age and context. The results of critical reviews have generally shown positive effects for CBT on anxiety and depression, although a number of limitations have

been identified in this research. These include the lack of active comparison groups, analysis of underlying mechanisms and long-term follow-up.

There is emerging evidence that universal school-based programmes can be as effective as targeted ones in reducing and preventing anxiety, and it is suggested that the EP is well placed to support the delivery of such interventions in schools.

2.5 'FRIENDS FOR LIFE'

Introduction

In this section the FRIENDS programme is introduced, which forms the basis of the intervention in the present study. Section 2.4 outlined the key features of CBT and presented evidence for its impact on anxiety and depression. Previous sections have described key theories relating to emotional distress and academic self-perceptions, highlighting the links between them, and evidence suggests these should be receptive to intervention. This chapter will focus more specifically on the components of FRIENDS, in order to consolidate the rationale for the researcher's hypotheses. A systematic review of studies involving the universal application of FRIENDS will then be presented to examine evidence for its impact in various contexts.

2.5.1 Theoretical Basis

The FRIENDS program was developed by Dr Paula Barrett as an intervention for the treatment and prevention of anxiety in children between 7 and 11. The programme has been widely researched and its impact acknowledged by the World Health Organisation, (2004). Its theoretical principles are based on the tripartite model involving the cognitive, physiological and learning processes that are seen to interact in the development, maintenance and experience of anxiety, (Barrett, 2004; see Appendix 2c). Further details about the rationale and structure of FRIENDS will be provided in Chapter Three.

2.5.2 Systematic Review of the FRIENDS Literature

A comprehensive, systematic search was conducted to identify studies that evaluated the effectiveness of the FRIENDS programme when delivered as a universal intervention. The purpose of conducting a systematic review was to scrutinize the protocols associated with previous FRIENDS research; to synthesize what has been learned about its effectiveness thus far and to identify areas that have been highlighted for future consideration (Petticrew and Roberts, 2006). The question under scrutiny was:

Does participation in a universal FRIENDS intervention reduce 'symptoms' of anxiety and/or depression for children aged between 7 and 13?

The specific search strategy, including inclusion/exclusion criteria for papers and search terms can be found in Appendix 2b. ***Figure 2.3*** illustrates the progression of the search and how it was refined:

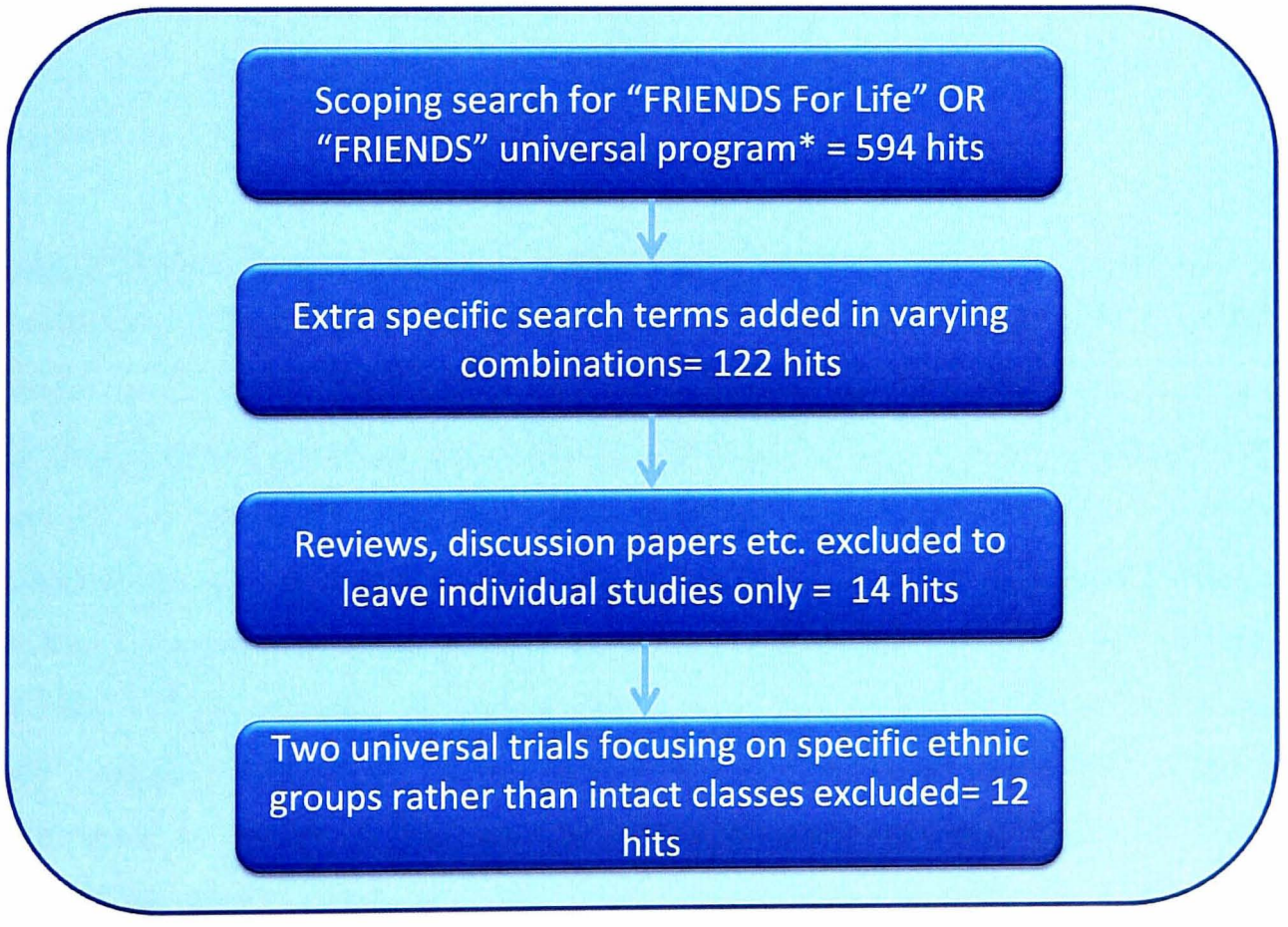


Figure 2.3 Flow chart to show systematic search procedure.

2.5.3 Description of studies

Twelve studies met the inclusion criteria outlined in Appendix 2b. Summary tables can be found in Appendix 2d and each will be summarised briefly before synthesizing the conclusions of this review.

Barrett and Turner (2001) conducted a randomised controlled trial (RCT) of a universal FRIENDS programme compared to a standard curriculum condition. 489 children (aged 10-12) participated from ten schools in Brisbane, Australia; the unit of randomisation in this study was the school. The authors also compared the effectiveness of teachers versus psychologists as group leaders. Outcome measures relied on self-report only and 88 children were lost in the final analysis. Results on the Spence Children's Anxiety Scale (SCAS; Spence, 1994) and the Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds and Richmond, 1978), showed a significant reduction in anxiety ($F(2,451) = 3.25$; $p<.05$ and $F(2,457) = 4.24$; $p<.05$ respectively), for both the teacher-led and psychologist-led intervention groups. High anxiety children in the intervention groups were more likely to move from the 'at risk' into the 'healthy' range but this sample lacked the power to detect statistical significance.

Lowry-Webster, Barrett and Dadds (2001) examined the therapeutic and preventive effects of a universal FRIENDS programme with 594 children (aged 10-13 years) from seven schools in Brisbane. Schools were matched for size and sociodemographics and randomly allocated to intervention (IG) or wait-list control (CG). Children were stratified at pre-test into high and low anxious groups according to their scores on the SCAS. At post-test, children in the IG reported fewer anxiety symptoms, regardless of their risk status, ($t(545) = 6.59$, $p<.05$). A greater percentage of children remained in the 'at risk' category in the CG; 75.3% 'at risk' in the IG showed significant benefits compared to 42.2% in the CG. No changes were reported in the RCMAS results for either group, and the Child Depression Inventory (CDI; Kovacs, 1981) revealed significantly lower scores for the high anxiety children in the IG

only. No attrition data is reported and the authors acknowledge the limitations of relying solely on self-report measures.

Lowry-Webster, Barrett and Lock (2003) conducted a one year follow-up to the above study. They found that the SCAS scores for the IG were lower than for the CG and that the high anxiety groups maintained their lower scores. Significant relationships between risk status and treatment group were found with 85% of the high anxiety and depression group 'diagnosis free' at twelve months compared to 31.2% of the CG. The authors concluded that intervention gains were largely maintained over twelve months, according to both self-reports and diagnostic interviews. The latter were only conducted at follow-up, however, so it was not possible to compare these scores with post-intervention, and the authors reported a poor response rate to the parent Child Behaviour Checklist (Achenbach and Edelbrock, 1991). 21% of the original sample was lost to attrition over twelve months.

Lock and Barrett (2003) carried out a longitudinal study evaluating the effects of FRIENDS at two developmental stages. The study involved a cohort of 737 children enrolled in Grade 6 (n=336) aged between 9 and 10 years, and Grade 9 (n=401) aged between 14 and 16 years. Participants were randomly allocated on a school basis to either a FRIENDS IG or a monitoring CG and completed measures of anxiety, depression and coping style. Participants were stratified into 'at risk' and 'healthy' groups, according to their results on the SCAS. Significant reductions in anxiety and depression were reported by the both groups at post-test, ($F(6,23)=45.49$, $p<.001$). The IG showed greater anxiety reductions at post-test and twelve month follow-up, however. Grade 6 reported significantly greater reductions than Grade 9 at 12 month follow up. No significant changes were found between the IG and CG in the 'at risk' group but the authors suggested that this may be due to attrition patterns.

Further follow-up measures to this study **(Barrett, Farrell, Ollendick and Dadds, 2006)** were taken at 24 and 36 months. 669 of the original sample completed the anxiety and depression measures again, with one school

withdrawing after twelve-month follow-up. The researchers found that intervention reductions in anxiety were maintained for students in Grade 6, with the IG reporting significantly greater reductions at long-term follow-up ($F(1,96 = 7.48; p<.01)$. No significant group differences were found at Grade 9 and the authors concluded that this supports Grade 6 as the more optimal time for intervention. A prevention effect was also demonstrated with significantly fewer 'high risk' students in the IG at 36 months compared to the CG. Strengths of this study include the RCT design and the large sample size, although it is difficult to control for confounding effects in the CG (including external referrals) over this time period and the study is also limited by the absence of diagnostic and multi-informant reports.

A parallel study carried out by **Barrett, Lock and Farrell (2005)** similarly examined differences between the effects of FRIENDS at two developmental stages. 692 participants from Grade 6 (n=293) and Grade 9 (n=399) were allocated to either a FRIENDS IG or a monitoring CG. Participants were stratified according to results on the SCAS at pre-intervention into 'low,' 'moderate' and 'high' risk groups, (measures were self-report only). Post-intervention results indicated significant reductions in anxiety ($F(2, 1.93)= 7.10; p< 0.001$) and depression ($F(2,1.97) = 5.37; p<.05$) across high and moderate risk groups, irrespective of intervention group status. Reductions were sustained at 12 month follow-up but the 'moderate' and 'high risk' groups in the IG showed significantly greater reductions than the CG, ($F(1,543)=7.29; p<.05$). At post-test, Grade 6 scores showed significant reductions compared with Grade 9, ($F(2, 1.93) = 13.066; p< .001$), and reductions were maintained at 12 month follow-up. The researchers concluded that overall, the findings supported universal interventions as potentially successful in reducing anxiety symptoms in children. However, 33% of the total sample was absent at both post-test and 12 month follow-up.

In 2005, **Stallard and colleagues** carried out the first published evaluation of FRIENDS in the UK. Their study involved 213 children aged 9-10 from 6 primary schools in the south-west of England. The intervention was delivered as a universal programme by trained school nurses. This study is severely

limited by its lack of control group and reliance on self-reports. However, it does reveal some useful trends as measures of both anxiety and self-esteem were taken. Post-test data revealed significantly lower rates of anxiety on the SCAS (t=2.950; p=.003) and significantly improved levels of self-esteem (t=2.950; p=.002) on the Culture Free Self-Esteem questionnaire (Battle, 1992). Post-test assessments for the 'high risk' group revealed a significant increase in self-esteem (t=4.789; p=.0001) and a significant decrease in anxiety (t=2.362; p=.023). By the end of the programme the status of 60% of children in the 'high risk' group positively changed. The authors concluded that FRIENDS appears to be an efficacious programme for promoting emotional resilience (reduced anxiety and improved self-esteem) in primary-aged children.

A further study **(Stallard et al, 2007)** replicated the original protocol with the addition of measures taken 6 months prior to intervention. The researchers found no significant change between Time 1 and Time 2 indicating that anxiety and self-esteem were initially stable. Post-test measures 3 months after programme completion revealed positive, significant changes in anxiety (F=5.84; p=.003) and self-esteem (F=2.98; p=.052) across time. Reduction in anxiety for the 'high anxiety' group were significant (F=5.30; p=.011) and the increase in self-esteem for the 'low self esteem' group was significant (F=5.78; p=.043).

A twelve month follow-up **(Stallard et al, 2008)** obtained data from 63 children (59% of the original cohort). Outcomes showed a significant effect over time for self-esteem ($F(3,323)$ = 6.55; p=.0001) and anxiety ($F(3,323)$=8.58; p=.0001). No significant differences were found between the Time 3 and Time 4 analyses indicating the long term maintenance of benefits. Of 9 children identified as 'high risk' at pre-test, 6 had moved into the 'low risk' category by 12months. No 'low risk' individuals moved into the 'high risk' range, indicating a preventive effect. Once again, the low sample size and absence of control group limit the generalisability of these results.

Three studies were identified that have evaluated the use of FRIENDS as a universal intervention in other countries. **Mostert and Loxton (2008)** explored the effectiveness of the FRIENDS programme in reducing anxiety symptoms amongst South African children. 66 children (aged 12 years, 30 girls and 36 boys) in two Grade 6 classes were recruited from a school near Cape Town, South Africa. The community was characterised as having low socio-economic status and most of the population were black, Afrikaans speaking. The researchers adopted a quasi-experimental, non-equivalent groups design with 4 and 6 month follow-up. Pre-intervention anxiety scores were high for both groups and the CG received the programme after Time 3, (4 month follow-up). Scores on the SCAS indicated a significant decrease in anxiety scores for the IG between Time 1 and Time 3 (pre-test and 4 month follow-up, p =.00) and Time 1 and Time 4 (6 month follow-up), but not between time 1 and Time 2 (p = .08), indicating little immediate impact but a cumulative effect in anxiety reduction over time. The CG's scores also declined across time, and showed an increased drop after starting the intervention, but reductions were not significant between any time points. Moreover, there were no significant between group differences at any time point. The authors tentatively concluded that their results supported the effectiveness of the programme over time, although they acknowledge the limitations of their study in terms of lack of randomisation, limited sample and exclusion of parental and booster sessions.

Gallegos (2008) conducted a large-scale evaluation of the Spanish version of FRIENDS (AMISTAD). 1,030 4th and 5th Grade students from 8 schools in a city in northern Mexico participated. The protocol constituted a quasi-experimental, non-equivalent groups design with 6-month follow-up. Participants were stratified according to anxiety levels and learning difficulties (LD) and allocated to one of four corresponding, non-overlapping groups. Schools were matched on socio-economic status and groups were matched on a range of measures at pre-test. Statistically significant improvements of a small magnitude were found for the overall sample and for children diagnosis-free and non-LD, in that those receiving the programme decreased the severity of their depressive symptoms and increased their proactive coping

skills. For children already showing risk of anxiety and/or LD, the programme did not produce any meaningful changes. However, children at risk of depression decreased by 2.6% in the IG and increased by 5.4% in CG, indicating a preventive effect. No significant increase in self-concept was found for children with LD, (the only group tested on this measure). Within the limitations of its protocol, the study appears methodologically robust and the author concluded that adaptations regarding culture, mode of delivery and content might be necessary to improve the programme's effectiveness in this context.

Rose, Miller and Martinez (2009) attempted to replicate the Australian findings in two Grade 4 classrooms in an urban school in western Canada. They employed a non-randomized control group pre-test, post-test design. A total of 52 students aged 8-9 years participated in the study. The class teacher delivered the programme over 8 sessions; no details of treatment integrity checks are provided. All children completed the Multidimensional Anxiety Scale for Children (MASC; March, 1997). Results indicated that all children reported reduced levels of anxiety post-test, regardless of group status and changes were not statistically significant for either group. The authors questioned the value of universal interventions in this kind of context where almost all children's anxiety scores fell within the normal range at the start. They also highlighted that although the MASC is sensitive to short term change, an effect might be seen subsequently as in several other studies. This study was limited by its small sample size and lack of multi-informant assessment.

2.5.4 Summary

Support for the effectiveness of universal FRIENDS programmes in reducing symptoms of emotional distress has been found in both Australia and the UK. However, studies in Australia have been predominantly carried out by the programme developers; the reliability of the UK studies has been compromised by the lack of control groups and some have been criticised

because the unit of analysis does not match the unit of randomisation (Briesch et al, 2010). Long term follow-up studies have generally suggested that treatment gains are maintained and preventive effects have been demonstrated by the lower rates of children in the intervention groups remaining or moving into 'high risk' categories when compared to controls. Evidence suggests that teachers and school nurses can be effective programme deliverers, as well as psychologists, although it has been found that the mean effect size for school staff implementers (*ES* = .22) is half that for specialist trained providers (*ES* = .56, Briesch et al, 2010). Some support has also been found for the benefits of intervening early with children at the primary stage, while several studies have shown a delayed effect with increasing impact after the completion of the intervention.

Studies conducted in other countries reveal more mixed results, however, and statistically significant effects for the IG at post-test have not been so apparent. Authors have pointed to reasons such as a mismatch between the programme content and the particular culture under investigation or the low anxiety rates at pre-test which calls into question the value of universal interventions.

Limitations of universal studies include the impossibility of blinding teachers to the status of the groups, introducing the likely confounding variable of increasing their attention and responsiveness to the dependent variables (Rose, Miller and Martinez, 2009). Moreover, in several studies, no measures of treatment integrity have been provided. Other problems include the unfeasibility of randomising individuals within a school context. This has led to some significant differences between groups at pre-test, while internal validity is threatened by the potential diffusion of treatments when intervention and control groups are within the same school. The studies so far have focused on limited geographical areas and in some cases, small sample sizes, which have reduced the statistical power of effects. Finally, many of the larger-scale studies have had significantly non-random attrition rates, (eg. Barrett et al, 2006), which may have affected the overall interpretation of results. It is apparent that further controlled studies are needed to address

some of the issues highlighted and to continue the important evaluation of this programme in the current climate of mental health promotion in UK schools.

2.6 Conclusion and Rationale for the Present Study

This study was conducted within the contemporary national context of mental health promotion in schools. It forms part of the Local Authority's evaluation of FRIENDS, which has been selected for dissemination through the countywide TaMHS project, as well as contributing data to the D & R Collaborative Programme in Educational Psychology.

This literature review has outlined the psychological theories and frameworks that support the potential of the FRIENDS intervention to reduce emotional distress and enhance academic self-perceptions through the development of coping cognitions and efficacy skills. Firstly, it has been shown that from a developmental perspective, scholastic anxiety becomes particularly salient in middle childhood, when children become increasingly aware of their achievements in relation to peers and expected norms, indicating that this might be a critical developmental stage to intervene to promote their sense of academic competence (section 2.2.2; Barrett et al, 2006). Secondly, it has been demonstrated how performance is significantly affected by an individual's sense of their own efficacy, as well as their skills; academic self-concept and self-efficacy have been presented as potentially crucial mediators in the development of academic proficiency (section 2.3.8). Robust evidence for the links between cognitive control, anxiety and self-efficacy has been presented (section 2.3.7.). Some support has been demonstrated for the developmental appropriateness and application of CBT programmes in alleviating psychological and emotional distress in children (section 2.4). The FRIENDS programme in particular has been shown to have some effectiveness as a universal intervention in this respect, although the evidence-base outside Australia remains limited (section 2.5).

2.6.1 A Unique Contribution

The present study, one of the first universal UK applications to include a control group, extends the previous literature on FRIENDS by exploring whether it is successful in reducing *and* preventing emotional distress as a whole-class intervention. This involves examining both the mean reduction in group scores for emotional distress *and* the number of children moving in and out of the elevated range of scores, ('preventive effect'). The study also aims to develop new understanding about the potential benefits of FRIENDS when delivered to all children, to see whether it has a positive impact on scholastic functioning by focusing on the variable of academic self-perceptions. Finally the study explores teacher-rated pupil behaviour to see whether any changes in externalizing problems correspond with variations in emotional distress.

The research questions are:

1. Does a class of Key Stage 2 children participating in a universal FRIENDS intervention report a significant reduction in emotional distress (ED) in comparison to a non-intervention control group?

2. Does a class of Key Stage 2 children participating in a universal FRIENDS intervention experience a preventive effect for ED in comparison to a non-intervention control?

3. Does a class of Key Stage 2 children participating in a universal FRIENDS intervention display significantly more positive academic self-perceptions than those in a non-intervention control group?

4. Does participation in a universal FRIENDS programme result in a significant improvement in teacher-rated pupil behaviour (reduced difficulties and increased prosocial scores) in comparison to a non-intervention control?

It is hypothesised that:

1. Experimental hypothesis: A class participating in a universal FRIENDS intervention will report a significant reduction in ED compared to a non-intervention control group.

Null hypothesis: There will be no significant difference in ED reported by a class attending a universal FRIENDS intervention and a non-intervention control group.

2. Experimental hypothesis: Reports from a class participating in a universal FRIENDS intervention will indicate a preventive effect for ED in comparison to those in a non-intervention control group.

Null hypothesis: There will be no difference in reported ED preventive effects between a class attending a universal FRIENDS intervention and a non-intervention control group.

3. Experimental hypothesis: A class participating in a universal FRIENDS intervention will report a significant improvement in academic self-perceptions in comparison to a non-intervention control group.

Null hypothesis: There will be no significant difference in academic self-perceptions reported by a class attending a universal FRIENDS intervention and a non-intervention control group.

4. Experimental hypothesis: Teachers of children participating in a universal FRIENDS intervention will report significantly improved behaviour (reduced difficulties and increased prosocial scores) in comparison to those in a non-intervention control group.

Null hypothesis: There will be no significant difference in teacher reports of behaviour (reduced difficulties and increased prosocial scores) for those

attending a universal FRIENDS intervention and those in a non-intervention control group.

The researcher also aims to address a number of broad concerns that have been highlighted through the literature review:

- Exploring the 'social' aspect of cognitive behavioural therapies as applied in universal contexts; in particular considering the interaction between 'internal' therapeutic techniques and the 'external' social environment (Bandura, 1986). This includes a consideration of the 'mechanisms' that might be exerting an influence (Kazdin, 2007).

- Evaluating the effects of a programme targeted at reducing emotional distress on aspects of cognitive competence (Grave and Blissett, 2004).

- Expanding the literature on the EP's unique role and contribution; exploring the effectiveness and practicality of supporting staff with the delivery of CBT (Rait et al, 2010).

Chapter Three describes how these research questions and aims have been operationalised in the present study.

3 CHAPTER THREE: METHODOLOGY

Introduction

In this chapter a number of theoretical positions and related methodologies that have been employed by social scientists are explored prior to explaining the present design and method. The researcher considers some of the contemporary debates around 'evidence-based practice' in educational psychology and argues for the value of adopting a critical realist perspective, in contrast to traditional scientific positivist paradigms. Section 3.5 onwards describes how this perspective influenced the procedures and measures that were employed to evaluate the programme's impact on the dependent variables. A key theme running throughout is the researcher's awareness of both contextual influences and her own contribution in shaping the project outcomes, and thus personal diary notes from observations and consultations are also reported. The chapter concludes with a consideration of the strengths and limitations of the present study.

3.1 The Quest for an Evidence Base

This study was undertaken amidst a political agenda that emphasised the development of 'evidence-based practice' with regard to mental health initiatives in schools (Wolpert et al, 2006). This approach is traditionally aligned with the field of medicine and upholds a hierarchy of evidence, with experimental paradigms, (particularly randomized controlled trials, RCTs), as the 'gold standard' in applied practice (Frederickson, 2002). The uncritical adoption of this protocol in the social sciences has been questioned, however, for a number of reasons: firstly, because the complexity of poorly controlled, 'messy' situations in the real world may be disregarded (Robson, 2002); secondly, it is proposed that the fallibility of scientific evidence is underestimated, while the often crucial contribution of researcher judgement and assessment is ignored (Hammersley, 2005); and thirdly, it is argued that

demonstrating that an intervention works reveals little about how it works and under what circumstances (Hughes, 2000). These debates have led to the conclusion that reliable evidence may derive from a spectrum of paradigms and data types and ultimately, the approach adopted should reflect judgements about the research question posed and the implications for practice in different contexts, (Hammersley, 2005; Ramchandani, Joughin and Zwi, 2001).

In order to discuss how these issues relate to the development of the present research protocol, it is necessary to critique some of the theoretical assumptions underlying these approaches and how these relate to models of current EP practice.

3.2 Ontology and Epistemology

Ontology refers to the branch of metaphysics that deals with the nature of being, (Oxford English Dictionary, 2002). The investigation of social phenomena necessitates consideration of whether reality exists as an external, objective entity or whether it is a product of individual, subjective consciousness (Cohen, Manion and Morrison, 2007). A related concern involves assumptions about epistemology, that is the nature and form of knowledge and how it can be reliably acquired and communicated (ibid, *p7*). As Cohen et al (2007) explain, the view that knowledge is "hard, objective and tangible" demands a quite different theoretical and methodological stance to that which views knowledge as "personal, subjective and unique," (*p7*).

3.2.1 Positivism

The fundamental supposition of positivism is that a single reality exists and it is the researcher's aim to discover the nature of that reality; a position that is aligned with the traditional experimental approach as applied in the natural sciences (Mertens, 2002). The assumptions of positivism include: the quest for objective knowledge gained from direct experience or observation; the

search for universal, causal laws that are founded on empirical regularities; hypothesis-testing and the gathering of quantitative data derived from strict rules and procedures (Robson, 2002). Although the positivist approach has advantages in terms of being explicit and systematic in the identification of relationships, (Robson, 2002), it has been subject to criticism when applied to the social sciences (Cohen et al, 2007). As Cohen et al (2007) describe, not only does its emphasis on the mechanistic and reductionist view of nature deny the complexity of inner experience, but its inclination towards a passive view of human behaviour excludes notions of choice, freedom, individuality and moral responsibility. Hammersley (2005) questions the validity of finding "simple causal relations" in educational contexts, due to the lack of standardisation of treatments, unreliability of outcome measures and the dynamic, adaptive nature of pupil and teacher behaviour, (*p90*).

3.2.2 Relativism

Conversely, relativism postulates that there is no external reality, independent of human consciousness (Robson, 2002). According to this view, what we perceive and experience is not a direct reflection of environmental conditions, but is socially constructed and mediated through history, culture and language (Willig, 2001). The associated epistemological approach suggests that diverse perspectives on the world should be explored, rather than evaluated in terms of their predictive power or explanatory value (Robson, 2002). Critics of this interpretative view have argued that subjective reports may be incomplete or misleading, while the focus on 'micro-sociological perspectives' inhibits the discovery of potentially useful generalizations about human behaviour (Cohen et al, 2007).

3.2.3 Post-Positivism

According to Mertens (2010), post-positivists concur with the positivist stance that a reality does exist, but due to human limitations, this can only be discovered within a certain realm of probability. Post-positivists strive for objectivity, but acknowledge that the theories, hypotheses, background

knowledge and values of the researcher can influence what is observed (Reichardt and Rallis, 1994). The researcher aims to remain neutral to prevent possible biases influencing their work and thus prescribed procedures are followed rigorously; associated methodologies reflect scientific paradigms but post-positivists acknowledge the difficulty of applying these with people (Mertens, 2010).

3.2.4 Critical Realism

While the 'naïve realism' of the positivist approach has attracted strong criticism, later realist interpretations retained the interest in causal relationships but de-emphasised the need for establishing regularities between them (Sayer, 1992). According to Robson (2002), critical realism can provide a model of scientific explanation that avoids both positivism and relativism. From this perspective, an independent reality does exist, but 'facts' are open to dispute; the task of science is to invent theories about the real world, which is multi-layered and complex (Robson, 2002). Realists view causal relations as tendencies, grounded in the context-specific interactions of generative mechanisms, (Outhwaite, 1998, see ***Figure 3.1***) and thus the question of "What produces the greatest change?" is rephrased as "What works best for whom, under what circumstances?" (Robson, 2002). The job of the researcher is to "manipulate the entire experimental system, so as to *manufacture* the desired interrelationship between independent and dependent variable" (Pawson and Tilley, 1997, *p60*, emphasis in original).

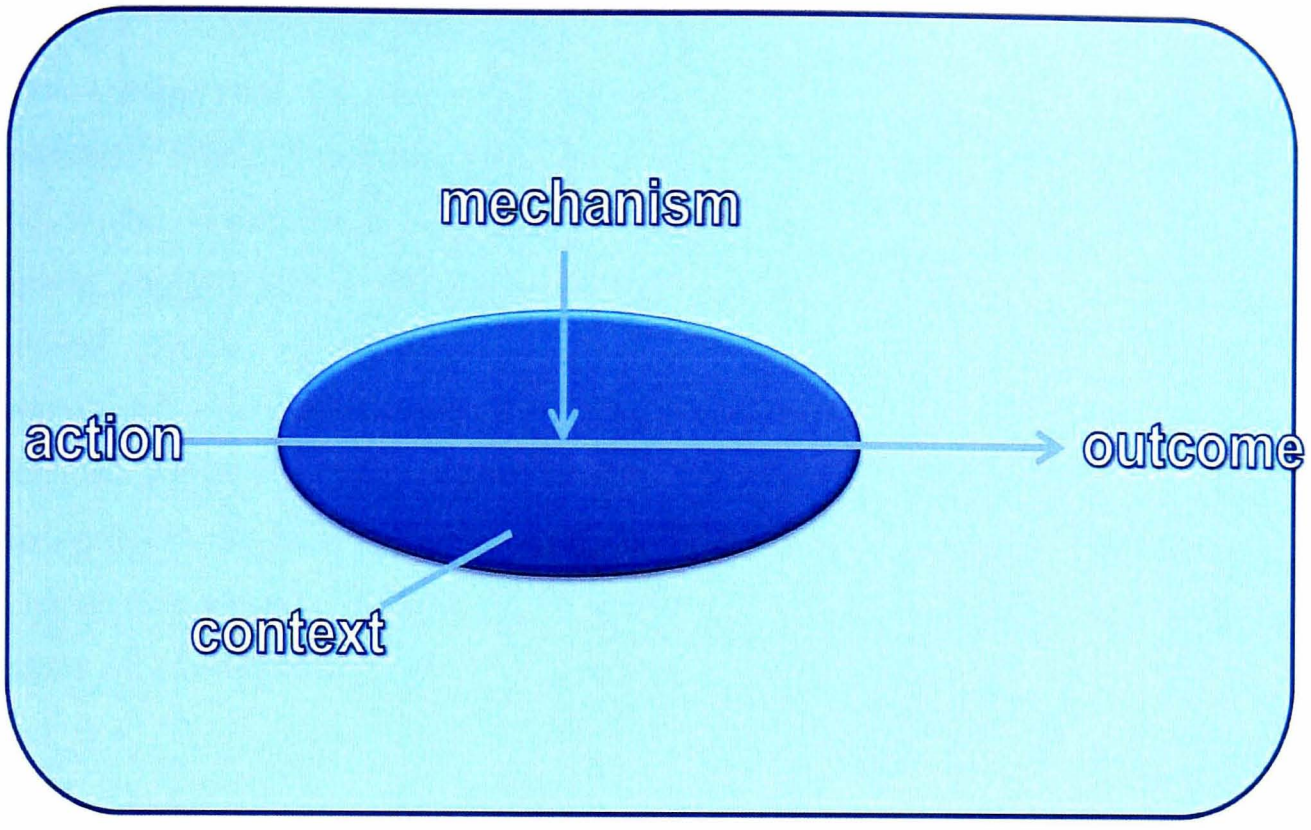


Figure 3.1: Representation of a realist explanation

3.3 Ontology and Epistemology in the Present Study

The present study constitutes an evaluation of a universal CBT programme. A leading theme of the evidence-based practice movement has been the imperative of "trying to do more good than harm," (Chalmers, 2003). Chalmers (2003, *p22*) suggested that this should be informed by "rigorous, transparent, up-to-date evaluations", although as has been apparent throughout this chapter, achieving this through traditional scientific practice in education is questionable (Hammersley, 2005). EP research has thus pointed to the value of integrating epistemological positions and methodologies, paying attention to both the outcomes of interventions and their processes and content (Frederickson, 2002; Miller and Todd, 2002).

The present study incorporates elements of the post-positivist paradigm in its attempt to establish causal relationships between variables. By extension, it

draws on critical realist philosophy in the consideration of mechanisms and acknowledgement of contextual influences upon its outcomes. As the researcher has been extensively involved in the programme implementation and in the manipulation of the experimental system, her position is more closely aligned with a critical realist perspective than the more detached, 'neutral' position advocated by post-positivism. The realist stance also permits the researcher to acknowledge the importance of reflexivity in the analysis process. This involves the researcher reflecting on her own standpoint in relation to the phenomenon being studied and identifying ways in which this shapes the research process and findings, (Willig, 2001). These issues will be expanded further in the discussion of design.

3.4 Designs

This section explores the methodological designs associated with the above paradigms as a precursor to explaining the present design.

3.4.1 Fixed Designs

Allied to the positivist and post-positivist paradigms, fixed designs are concerned with aggregates, group properties and general tendencies, (Robson, 2002). The features of a true experiment can be summarised as manipulating an independent variable and measuring its effect on a dependent variable, while holding other variables constant (Coolican, 2009).

Randomized Controlled Trials (RCTs)

In a RCT, participants are randomised to an experimental (intervention) group or to a control group that does not receive any particular treatment. The purpose is to control for confounding variables and to strengthen conclusions regarding intervention effects, although Coolican (2009) emphasised the importance of maintaining a critical awareness when reporting causal links. Advantages of experimental designs are taken to be establishing high validity (the accuracy of the result); reliability (the stability of the result) and

generalisability (the extent to which the results might apply across contexts, populations and times; Robson, 2002). Researchers in the post-positivist paradigm indicate that experiments are the most compelling method of establishing causation and this is particularly important for the evaluation of educational innovations (Moore and McCabe, 1993; Slavin, 2002).

As highlighted previously, however, establishing definitive causal relationships with social phenomena is problematic, perhaps even undesirable. Reactivity effects, ethical issues, the unfeasibility of random assignment, issues of validity surrounding the integrity of treatment, the reliability of outcome measures and the lack of control over extraneous variables are just some of the issues highlighted by researchers, (Robson, 2002; Coolican, 2009). Moreover, Mertens (2010) suggests that the attempt to control variables (such as the background characteristics of participants) may actually be misleading in real world settings because it risks oversimplifying and distorting how social phenomena occur. Meanwhile, Maxwell (2004) argues against privileging RCTs as the "gold standard" in educational and psychological research, suggesting that a realist view of causality can legitimately be explored through qualitative approaches, (*p3*).

Quasi-experiments

Quasi-experiments include comparisons between different levels of a treatment variable but do not involve random allocation, (Mark, 2010). As Cohen et al (2007) pointed out, in educational research it is often not possible for investigators to undertake true experiments; thus, quasi-experimental approaches have been termed "compromise designs" where random assignment is impractical, for investigators working in schools, for example (*p282*). It has been suggested, however, that quasi-experimental models involving single groups or post-test only designs do not permit the researcher to reliably infer whether any difference in group/test results can be attributed to the treatment (Robson, 2002). Indeed, authors concur that quasi-experimental designs are more vulnerable than RCTs in terms of their 'threats to validity' (Cohen et al, 2007; Reichardt and Mark, 2001; see Section 3.16 for

a full description in relation to the present study). Conversely, as Shadish, Cook and Campbell (2002) argue, the retention of intact groups may have advantages in applied research as participants' psychological or social responses to treatment may be affected by randomisation processes.

3.4.2 Flexible Designs

Associated with the relativist paradigm, flexible designs are usually concerned with the collection and interpretation of qualitative data through interviews and observations, for example. The focus is upon evolving design, the presentation of multiple realities, and emphasis upon participants' views (Robson, 2002, *p166*), the assumption being that data, interpretations and outcomes are rooted in particular contexts, (Mertens, 2010, *p19*). Within this design, the researcher adopts a reflexive position to identify potential sources of their own bias (Ahern, 1999, cited in Robson, 2002). Flexible designs may elicit valuable information about social phenomena by unravelling the individual meanings and personal constructs that people attach to their experience and how these relate to their behaviour (Mertens, 2010), but entail careful consideration of descriptive, interpretive and theoretical validity, for example, (Maxwell, 1992).

3.4.3 Evaluation Research

The current emphasis upon 'accountability' in education has encouraged a trend towards evaluative research, which assesses the effectiveness of given policies and projects (Robson, 2002; Cohen et al, 2007). It is argued that evaluations are inextricably bound with the concerns of stakeholders and the process of change, with findings influencing future development and implementation (Robson, 2002; Mertens, 2010). Evidence suggests that contemporary evaluation models are increasingly incorporating realist perspectives that move away from 'hierarchies' of research designs equated with strength of evidence, in preference for methods that combine effectiveness data with additional information about the context and process of service delivery, (Petticrew and Roberts, 2003; Chatterji, 2008). The

present study, drawing upon both post-positivist and realist paradigms, offers a summative appraisal of the outcomes of the FRIENDS programme but through the research diary, seeks to evaluate some of the processes that might be operating within its particular context.

3.5 Rationale for the Present Design

This study evaluates the impact of FRIENDS within a given community context. Stakeholders require information about the impact of the programme on reducing and preventing emotional distress, improving academic self-perceptions and improving pupil behaviour. Shadish, Cook and Campbell (2002) describe how contemporary models of evaluation are tending towards realist perspectives that integrate experimental and qualitative approaches through an iterative process. Whilst fully acknowledging the value of this stance, the researcher considered that the scope of this study, in addressing two extensive theoretical themes, (anxiety and academic self-perceptions), would allow only for a detailed analysis employing one of the approaches. Quantitative analysis has thus been selected as an efficient method for assessing change across a number of participants (Frederickson, 2002), although the potential impact of contextual factors has also been documented.

3.6 Procedure

3.6.1 Design

This study comprised a quasi-experimental pre-test, post-test, non-equivalent two group design. The intervention ran in a Year 5 class with the parallel class acting as the non-intervention wait-list control. A quasi-experimental approach was chosen over a randomised controlled trial because it was considered in consultation with project school staff that the latter might have a detrimental effect on the children's behaviour. The researcher was also keen to promote the application of the programme strategies throughout the school

day to encourage the generalisation of skills. It was therefore considered that keeping the classes separate was an important factor in minimising diffusion of treatments (Cook and Campbell, 1979).

3.6.2 Data Collection Timeline

The independent variable in this study was participation in the FRIENDS programme. The dependent variables were pupil and teacher ratings of emotional distress, self-reported academic self-perceptions and teacher ratings of pupil behaviour. Measures (described in section 3.10) were taken at two time points, the first prior to the commencement of the intervention and the second following the experimental group's completion of the programme and prior to the control group's participation. A third administration of the emotional distress measure is planned for the control group during the Summer Term 2011 for ethical reasons, although this will not be reported in the present study. (See section 3.14). ***Figure 3.2*** illustrates the structure of the design and Appendix 3d shows the study timeline.

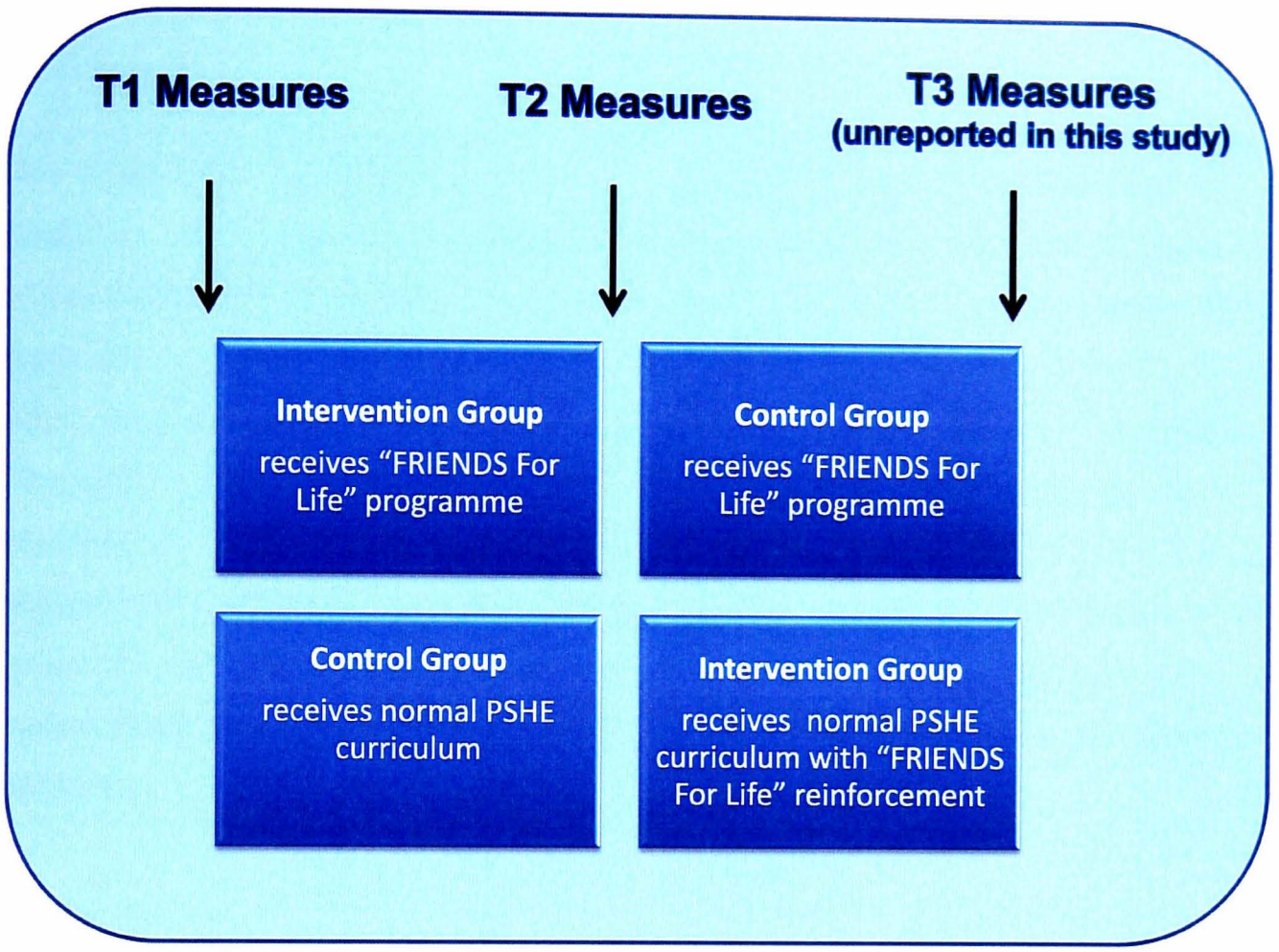


Figure 3.2: Diagram to show the intervention/control group design and measures

Although the predominant design model is fixed, the researcher was actively involved in supporting school staff to deliver the FRIENDS programme, and it was considered important to reflect upon the impact of this model of delivery. The researcher kept a log to record her visits, together with any relevant information about how the programme was being implemented. This involved making notes of the consultations with the group leader about adaptations to the basic programme, mode of delivery and strategies to reinforce the taught skills, for example.

3.6.3 Context

The researcher is employed in a local authority in a town in the East of England. A number of schools participating in the TaMHS project were considered but rejected, due to them already having a number of measures to complete. Other schools were approached directly and via colleagues, but while several showed interest in delivering the programme as a targeted intervention, many were reluctant to adopt it universally due to already well-established PSHE programmes. The researcher was also restricted to schools with two-form entry in order to obtain a comparable control group. Reichardt and Mark (2001) point out that selection differences in quasi-experiments can be minimized by recruiting individuals from the same institution.

The researcher thus proceeded to approach her own link schools and the two that fulfilled the organisational requirements agreed provisionally to participate. The researcher recruited the support of school staff and senior management through a short presentation detailing the aims of the project and sought permission from the head teachers (See Appendix 3a). However, one of the schools withdrew just prior to commencement of the project, substantially reducing the data set. The researcher's focus thus shifted from concentrating purely on outcomes, to embracing a more detailed exploration of how the intervention impacted upon one particular classroom context. This permitted consideration of how specific adaptations and processes might operate to produce the given outcomes (Kazdin and Nock, 2003). The researcher conducted a pilot study with a Year 3 class in the project school during the Summer Term 2010, (see Section 3.15).

The project school, located in an area of relative socio-economic deprivation within the town, agreed to participate due to an increasing interest in promoting mental health initiatives. According to the most recent Local Authority statistics available, 22% of the population in this geographical area was from an ethnic group other than White British and a large proportion were

within the 50% and above 'most deprived' category. 25.8% of children in this area were considered to be living in poverty and the number of first time entrants into the youth justice system was above the national average. Average point scores for Key Stages 1, 2 and 4 were well below the national average.

3.6.4 Participants

The project school had 378 children on roll. 58% of the population was White British and the remainder from a variety of Black, Asian, mixed and other ethnic backgrounds. 50% were eligible for free school meals.

Participants were drawn from Year 5 as staff had identified this year group as being particularly emotionally vulnerable. The class chosen to receive the intervention first had children with a history of challenging behaviour, according to staff report. This age group was also chosen because it matched the standardisation requirements for the researcher's measures.

42 participants aged 9-10 years were initially recruited across the two classes but two children from each class left the school during the project and their data was subsequently discounted from the analysis. ***Table 3.1*** illustrates the (adjusted) demographics for each group.

	Intervention	Control
Total	18 (7 males, 11 females)	20 (8 males, 12 females)
English as an Additional Language (EAL)	6	5
School Action or above on Special Educational Needs (SEN) register	4	10
Additional support during past academic year (eg. Literacy/numeracy; mentoring; therapeutic group work)	14	17
Working at Level 2 or below in reading	6	9

Table 3.1: Demographic information for the intervention and control groups

3.6.5 Stakeholders

According to Mertens (2010), identifying stakeholders permits the researcher to ascertain the breadth of the impact of the evaluation.

Stakeholders in the present study include:

Local Authority: endorsed the implementation of FRIENDS across the county and TaMHS data revealed positive results; outcomes may influence future recommendations, especially as the Educational Psychology Service comes under increasing pressure to generate its own income through training.

University of Nottingham: contributes to the National D&R programme; has the potential to influence the national agenda in research and directions for future EP practice.

Project School: has a vested interest in terms of time and resources; outcomes may influence future decision-making about the implementation of this and other therapeutic programmes in conjunction with EP support.

Parents: have an interest in the assessment and well-being of their children; require information about the overall effectiveness of the programme and any concerns that arise from individual measures.

Children: have an interest in terms of potential changes to their emotional well-being.

3.7 Intervention

3.7.1 Rationale

Full details of the FRIENDS programme can be found in the Group Leader's Manual (Barrett, 2004); a brief summary of its philosophy and components will be included here.

The FRIENDS programme was designed to assist children aged 7-11 years in developing life-skills to effectively cope with difficult and/or anxiety-provoking situations. According to the author, the programme aims to normalise the emotional state of anxiety, build emotional resilience and promote self-confidence and problem-solving abilities. Barrett (2004) asserts that the programme captures the essence of Australian culture and is aimed at an appropriate developmental level. The acronym helps children to remember the CBT principles and skills:

F = Feelings

R = Remember to Relax

I = I can do it, I can try my best

E = Exploring Solutions and Coping Step Plans

N = Now reward yourself

D = Don't forget to practise

S = Smile, stay calm for life

3.7.2 Teaching Philosophy

The programme is founded on two specific teaching philosophies:

Peer Learning: designed to be implemented in a naturalistic environment, involving a group of participants of the same age; encourages individuals to observe and help each other; promotes learning in context with peers, providing opportunities for participants to practise new skills in a safe environment.

Experiential Learning: encourages participants to learn from their own experience; encourages them to play an active role in generating ideas; emphasises that group leaders and participants have valuable knowledge and experiences to bring to the group.

3.7.3 Training

It is recommended that group leaders become familiar with the principles, skills and techniques offered through special Group Leader training "so they are fully aware of the possibilities and limitations of their role." Accredited FRIENDS training is organised through the Pathways Health and Research Centre or its approved partners (www.friendsinfo.net).

3.7.4 Structure

The programme consists of 10 sessions, each running for approximately 45-60 minutes. It is suggested that the sessions are run weekly for maximum effectiveness.

Outline of sessions:

1. Introduction to the Group.
2. Introduction to feelings and their association with behaviour.
3. Feelings, physiological symptoms of worry; relaxation.
4. Identifying self-talk; red thoughts and green thoughts.
5. Attention training; exploring solutions and coping step plans.
6. Problem-solving skills; coping role models; social support plans.
7. Rewards.
8. Practising the FRIENDS skills.
9. Generalising skills to various difficult situations.
10. Skills for maintenance of the FRIENDS strategies; preparing for minor setbacks.

In addition to the manual, there are individual children's workbooks to complement the programme. Two booster sessions should be delivered one month and three months after the completion of the programme and there are materials for four optional parent sessions.

According to the manual, a single group leader is most effective taking small groups of no more than 12 participants. It is recommended that for larger groups run in the classroom, more than one group leader be involved in running the programme. The author suggests that a high ratio of leaders to participants increases the attention that may be given to each individual and encourages interest and involvement in the activities.

3.8 Implementation of FRIENDS in the Present Study

3.8.1 Type and Duration of Intervention

The programme was implemented as a universal intervention involving all children in the class. All ten sessions of the programme were delivered on a weekly basis, between October 2010 and January 2011. A few weeks were omitted due to holidays or other school commitments. The project took place on Friday afternoons and the average length of each session was around 90 minutes.

Parental permission was actively sought for each child via a letter, following the opportunity to attend a short meeting to describe the overall aims of the programme and to introduce the study (see Appendices 3b and 3c). This meeting was very poorly attended and it was necessary to engage the help of the school's family support worker to reach all of the parents over a number of weeks. Due to the reported and observed difficulty of engaging parents generally, a decision was taken by the researcher and group leader not to include the parent component of the programme on this occasion.

3.8.2 Control Group

The parallel class in Year 5 acted as the wait-list control group (CG) who were to receive the intervention during the Spring Term 2011. They were located in the adjacent classroom and some of the members were taught in the Intervention Group (IG) classroom for literacy and numeracy. The CG had two class teachers who job-shared throughout the week. While the IG participated in FRIENDS, the CG received their normal PSHE curriculum based on Social and Emotional Aspects of Learning (SEAL, DCSF, 2005). The CG had access to the same additional interventions (such as Relax Kids, www.relaxkids.com) as the IG. The group leader reported that as far as was possible within the school environment, the CG had not been exposed to the specific components of the intervention. Reports from the CG class teacher

and the children's naive enquiry about the programme to the researcher during the T2 measures supported this observation.

3.8.3 Programme Implementers

The programme was organised and implemented by the Learning Mentor, with the support of the researcher, who attended 6 out of the 10 sessions. The original arrangement was for the Learning Mentor to deliver independently with the support of the (untrained) class teacher. However, two sessions into the programme, the group leader felt that successful implementation depended upon a higher adult to child ratio, with the additional psychological expertise that had been integrated during the pilot phase (see section 3.15).

The group leader was in her 20s, white British and had background qualifications and experience in delivering therapeutic group work and understanding and managing children with social, emotional and behavioural difficulties. Both she and the researcher had attended the accredited Pathways training delivered by the county Educational Psychology Team. The Year 5 class teacher and learning support assistant also occasionally supported the delivery of the programme, although they had not received the training. The group leader had ongoing access to consultation with the researcher and regular discussions were held following the weekly sessions to discuss practical issues of implementation, the children's receptiveness to the programme and their generalisation of skills in the intervening week. She later reported that her confidence in delivering the programme and adapting it to the children's levels and needs improved over the course of the intervention, and this was largely due to having established a good relationship with the class (refer to discussion of 'therapeutic alliance' in Chapter Five).

3.8.4 Treatment Integrity

The group leader and researcher completed the FRIENDS Treatment Integrity Scales (Barrett, Lowry-Webster and Turner, 1999), for three of the ten sessions. Both rated that the aims for each activity had been achieved either 'extremely well' or 'moderately well' and none as 'not very well' or 'not at all', indicating reliable treatment fidelity. It is important to stress that the authors condone some freedom and creativity in the way that activities are presented, asserting that the scales assess whether the group leader has met the objectives for each activity rather than their strict adherence to the format of each one. Treatment integrity could have been strengthened further, however, by including an independent rater.

3.8.5 Adaptations to the Basic Programme

The Group Leader's Manual was used as a basis for designing weekly lesson plans. Through consultation with school staff, it was decided that certain sections required amending to make them more culturally applicable to this sample, and thus some of the Australian terminology was substituted, (eg. 'yard' for 'drive' and 'principal' for 'head teacher'). The researcher and group leader also discussed on a regular basis whether the mode of delivery suggested was suitable for this sample of children. For example, it was considered that the official Children's Workbooks involved large amounts of text that some might find intimidating and the decision was taken to have individual pupil scrapbooks for the children to record their responses instead. However, the tasks in the books were adhered to as far as possible with the group leader providing her own worksheets and writing frames and completion of the Treatment Integrity Scales indicated that key learning objectives were covered adequately each week (see previous paragraph).

Further adaptations were made to reinforce specific aspects of the programme for this sample of children. For example, it was noted in joint consultation that the FRIENDS coverage of the relaxation component was

limited and it was agreed from observations that these skills needed continual reinforcement. The group leader therefore integrated some of the 'Relax Kids' materials at the start or end of some sessions to offer additional practice of relaxation techniques. Finally, the children were explicitly encouraged to relate the strategies taught (particularly the 'green thoughts' and 'coping step plans') to learning and social problems they experienced in school, such as revising for a spelling test, facing a difficult maths problem, public speaking or resolving a dispute on the playground.

3.8.6 Additional Support

In addition to the programme sessions, the group leader offered a homework club during lunchtime to encourage those children who received little support at home to complete the weekly set tasks. This opportunity was taken up widely and increased as the intervention progressed, (50%+ by week 4). Children were rewarded with praise and stickers for the completion of extra work. Much of this work was deliberately centred upon encouraging the children to apply the skills and techniques taught to problems they encountered in school. The group leader also made herself available during lunchtimes for mentoring children who had individual issues arising from the programme. This gave them the opportunity to talk through concerns or to reinforce the application of taught skills. In addition to this, the group leader and the researcher set weekly 'challenges' to encourage the children to practise and generalise their learning. For example, children who were able to change a 'red thought' to a 'green thought' in relation to a difficult piece of work or an incident on the playground were to report their success to the group leader in return for a small reward. A small, but again increasing proportion of the class responded to this ongoing incentive.

3.8.7 Classroom Context

Classroom displays were used to reinforce some of the key concepts such as the FRIENDS acronym, the 'green thoughts' and 'coping step plan.' The group leader also displayed a 'feelings ladder' on the wall, which the children

were encouraged to position themselves on regularly. Where possible, the group leader followed up individual children who had evaluated their feelings negatively and offered support if required. In accordance with the ethical arrangements described in section 3.14, a 'worry box' was posted for the children to share concerns that they did not wish to approach an adult about directly and the group leader or teacher addressed these anxieties individually. Staff reported that the worry box was used consistently throughout the intervention but notably fewer concerns were received towards the end.

According to the group leader's report, the class teacher had been observed to reinforce the taught strategies throughout the week and to encourage the children to apply them to difficulties in school.

3.9 Researcher Reflexivity

Robson (2002, *p173*) summarises a number of areas in which the researcher can bias outcomes and interpretations. Although these are presented in relation to flexible designs, they have guided the researcher's thinking in the following analysis. The purpose is to illustrate how some of the researcher's values and motivations may have influenced the implementation process.

The researcher acknowledges the integral role that she played in the implementation of this project and how this contributed to 'manipulating' the experimental environment. Her stance as an educational psychology doctoral student privileges the promotion of psychological interventions and this undoubtedly influenced the 'neutrality' of this research. The desire to demonstrate effectiveness for the stakeholders incited the researcher to intensively promote the elements of the programme both within and outside the taught sessions; for example, by intervening during lessons to reinforce the CBT model and coping strategies and encouraging staff to promote them throughout the week. She also placed value on building a relationship with the children, giving them abundant praise and positive feedback for

contributing ideas and showing effort in engaging in the activities. It was noted by the class teacher and group leader that the children received considerably more individual verbal praise and stickers during FRIENDS sessions than they did for other curricular lessons. The researcher also offered extra small prizes for those who could report the generalisation of their skills to the group leader throughout the week.

The researcher was aware of her desire to make the project an enjoyable and worthwhile experience for the school. She therefore spent considerable time building a rapport with school staff and in weekly debriefing sessions with the group leader. Teachers were given small tokens of appreciation for completing their questionnaires.

Finally, the researcher's decision not to include the parent sessions was influenced by the difficulty in getting families involved. Staff reported that this was an historical problem, despite a variety of incentives being provided in the past. Although parents were invited to a session with refreshments and a raffle, to introduce the programme and to obtain consent, only around five chose to attend. Following a cost-benefit analysis of attempting to recruit parents for the special sessions and measures, it was predicted that this would not have yielded adequate responses. Thus the attempt was abandoned, although it is acknowledged that this is a major drawback of the overall implementation and analysis.

3.10 Pupil Outcomes

This section outlines the measures that were chosen to assess changes in emotional distress, academic self-perceptions and teacher ratings of behaviour. Normative measures have been selected here to describe group profiles and whilst the PI-ED and SDQ, for example, are used in clinical, individual assessment contexts, this is not their purpose here. This point will be referred to in Chapter Five, particularly in relation to the clinical cut-off points for the PI-ED.

3.10.1 The Paeditaric Index of Emotional Distress (PI-ED; O'Connor et al, 2010). [GL Assessment, 2010] (See Appendix 3e)

Rationale for choice

The PI-ED is a recently published self-report measure that screens children aged 8-16 for 'symptoms' of emotional distress. Developed from the Hospital Anxiety and Depression Scale, (HADS, Zigmond and Snaith, 1983), which is reported as a valid and reliable means of detecting anxiety/depression in adults, the PI-ED is targeted at paediatric populations in both mental health settings and schools. The developers suggest that it can be administered when there are specific concerns about a child's emotional well-being or to screen a general population as an index of therapeutic change. The measure came to the notice of the researcher via her supervisor at the University of Nottingham.

In selecting an appropriate measure of anxiety/depression for this study, a number of scales used in previous FRIENDS research were evaluated (eg. the Spence Children's Anxiety Scale; Spence, 1997 and the Multi-dimensional Anxiety Scale for Children; March, 1997), but rejected for various reasons. These included their length and emphasis on diagnostic categories, which the researcher considered to be incongruous with the dimensional view of mental health discussed in Section 2.1. Although the PI-ED is presented as a clinical measure that may serve as an index of clinical change, the developers also emphasise its value in assessing general levels of emotional distress for research purposes. As the intention of this study was to measure *group changes* in anxiety/depression, the researcher considered that the PI-ED constituted the 'best fit' option of those scrutinised. A brief measure, promoted as using language and concepts that are easily accessible to children and having a reading age of seven years; it fulfilled the aim of producing standardised scores for reliable comparisons in a universal research context, whilst de-emphasising the notion of individual categorical 'disorders.' The identified cut-off points for "clinically significant levels of emotional distress" (O'Connor et al, 2010, *p9*) were viewed critically in this

study and applied as thresholds for tracking dimensional change, as opposed to signalling a 'clinical category' of emotional distress. However, for ethical reasons, children scoring within the higher range on the measure were brought to the attention of the group leader for further consultation. [See Section 4.1.2. for further discussion of the application of clinical cut-off points].

The publishers granted permission for the researcher to use the PI-ED prior to its general distribution, by requesting a short case study to be submitted following completion of the project.

Standardisation Method

The PI-ED was standardised on an initial sample of 1108 participants from 27 schools from Ayrshire and Arran in Scotland and Nottingham City in England. In total, 47% of the sample was female (n=521) and 89% of the sample reported their ethnicity as White UK. The age range of respondents was 7-17 years (mean age =11.93 years; standard deviation= 2.33). The PI-ED was presented alongside the Beck Youth Inventories (Beck et al, 2005), which are reported as a widely used, valid and reliable measure within this population.

Diagnostic Sensitivity and Test-retest Reliability Method

This was determined on a population of paediatric outpatients (n= 117) aged 8 – 16. The PI-ED was tested on two occasions together with the Beck Youth Inventories (Beck et al, 2005) and the Diagnostic Interview Schedule for Children (C-DISC, Schwab-Stone et al, 1996).

Reliability and Validity

The internal reliability of the scale was assessed using Cronbach's coefficient alpha. A scale is said to be reliable if the value of this coefficient is equal to or greater than 0.70 (Nunnally, 1978). The coefficient values across the school and clinical samples indicated that the cothymia ('emotional distress') factor and its comorbid symptoms of anxiety and depression were reliable. Test-

retest reliability for the clinical sample showed cothymia to have a test-retest correlation of 0.81 ($p<0.001$) anxiety of 0.71 ($p<0.001$) and depression of 0.77 ($p<0.001$), indicating stability over time.

Validity was assessed for the school sample by means of zero-order correlations and linear regression procedures against the Beck Youth Inventories. Results indicated PI-ED cothymia was associated with the Beck Anxiety Scale (BAI-Y) and the Beck Depression Scale (BDI-Y) equally; PI-ED anxiety was associated more strongly with BAI-Y and PI-ED depression with BDI-Y.

Age, gender and ethnicity bias

Investigations of systematic variation showed that PI-ED cothymia exhibited no bias with regard to age and ethnicity but girls reported significantly higher levels than boys.

Diagnostic Sensitivity Result

Comparison of sensitivity and specificity co-efficients for the clinical sample revealed a clinical cut-off value of value of 10 for boys and 11 for girls, although recent data collected by the authors suggests that this may be revised to at least 15 (Personal communication with E. Ferguson, 01.04.11).

3.10.2 Myself-As-Learner Scale (MALS, Burden, 1998b). [NFER Nelson, 1999]. (See Appendix 3f)

Rationale for Choice

For the purposes of tracking changes in academic self-perceptions the researcher required a scale that focused on the self as a learner. General measures of self-concept or self-efficacy were disregarded (eg. Self-Perception Scale for Children, Harter, 1985). The MALS presented as a reasonably brief measure that was easy to administer, complete and score and the author cites its foundations in self-concept and self-efficacy theory.

Many of its key elements reflect the constructs that the researcher hypothesised would be targeted through the intervention:

1. Enjoyment in problem solving
2. Confidence about schoolwork/academic self-efficacy
3. Confidence about learning/learning self-efficacy
4. Taking care with work/careful learning style
5. (Lack of) anxiety
6. Access to and use of vocabulary in problem-solving
7. Confidence in dealing with new work
8. Confidence in problem-solving ability
9. Verbal ability/fluency
10. Confidence in general ability (Burden, 1998a)

Standardisation and Reliability

The scale was standardised on a sample of 389 Year 7/8 pupils attending a large urban secondary comprehensive school. These produced a set of norms, which suggested that a score between 60 and 80 out of 100 (mean = 71, SD, 10.5) fell within the average range. An alpha reliability index of 0.85 was reported, indicating strong internal consistency. The MALS has also been tested against measures of cognitive ability and basic attainments in literacy and numeracy and moderate positive relationships were found. Concurrent validity measures were sought for the MALS against the Connell Children's Perception of Control Scale (Connell, 1985). Weak negative correlations were found for 'unknown cognitive control' and 'control by powerful others' and a weak positive correlation for 'internal cognitive control.'

In the publication manual, Burden (1999) indicated that data from a wider sample of schools and age ranges was being collected to investigate possible developmental trends and the effects of school/classroom context on pupil responses to the MALS. In an e-mail to the researcher on 29.01.10 (see Appendix 3g) Burden indicated that although this additional data had been scrutinised, it was as yet unpublished. He concluded that the scale's reliability

below the age of nine years was "suspect" and recommended that for younger age groups every question was read aloud to ensure that each child understood what was being asked of them and how to respond appropriately. The researcher acknowledged this advice by selecting Year 5 for the main study (as opposed to the original plan to use the pilot Year 3s) and ensured that recommendations were adhered to in the administration of the measure.

3.10.3 Strengths and Difficulties Questionnaire – Teacher version (SDQ, Goodman, 1997). [Freely available via www.sdqinfo.org]. (See Appendix 3h).

Rationale for Choice

The SDQ is a brief behavioural screening questionnaire about 3-16 year olds. The SDQ measures 25 psychological attributes, divided into 5 scales: emotional symptoms; conduct problems; hyperactivity/inattention; peer relationship problems and prosocial behaviour. The first 4 of these combine to produce a 'total difficulties' score. This data was included to complement the pupils' self-report measures; individual subscale scores would provide additional information about the variable of emotional distress and also permit an exploration of some of the externalising behaviours that may be associated with it (see Chapter Two, section 2.2.5).

The SDQ can be used for clinical assessment of mental health or behavioural difficulties, evaluating outcomes for interventions and research, epidemiology and screening.

Standardisation and Reliability

The SDQ has been standardised on a number of cross-cultural populations. Chronbach's alpha coefficients revealed satisfactory internal consistency (mean 0.73) with 0.80 for the total difficulties score. The reliability and validity of the measure have been supported by a number of cross-cultural community and clinical samples (Koskelainen, Sourander, & Kaljonen, 2001; Hawes and Dadds, 2004).

3.11 Reliability of Children's Self-Reports

It is suggested in contemporary paediatric health literature that children are reliable and accurate reporters of their own health status when assessment methodologies are sensitive to their developmental level and cognitive competencies (Limbers, Newman and Varni, 2008; Bevans and Forrest, 2009).

However, a body of research highlights the importance of applying caution when interpreting self-reports and the need to critically consider contextual influences on reliability (Coolican, 2002; Lewis and Lindsay, 2000). Potential issues include: the interaction between the child and research situation which may cause him/her to respond in a socially desirable way or feel inhibited or unable to express their views; the tendency of younger children to view themselves in an unrealistically positive way, and the complex trajectories of children's behaviour and performance which necessitate a creative approach to understanding "what it is that is changing in development" (Dockrell, Lewis and Lindsay, 2000, *p49*). Begley (2000) has countered that the aim of research is not necessarily to establish how 'precise' children's self-perceptions are in comparison to some 'objective' standard. "It is the children's conception of themselves that will affect their self-concept, regardless of how accurate their self-perceptions are," (*p109*). However, Wigelsworth and colleagues' (2010) observation that young children's reports tend to be biased towards the "here and now," rather than offering summative judgements over time, questions whether *typical* response measures, (such as the ones used in this study), are capable of capturing stable changes. (See Wigelsworth et al, 2010, for a review of the discrepancies between *typical* and *maximal* measures). Dockrell et al (2000) advocate the critical application of a range of triangulated methods, appropriate to the research question, age and characteristics of the child to strengthen the validity of self-report findings.

The limitations of the current measures and methodology will be highlighted in Chapter Five in light of these observations.

3.12 Administration of Measures

The PI-ED and MALS were administered to the IG and CG at two time points (T1 and T2). T1 measures for each class were taken in September 2010, prior to the commencement of the programme, and were conducted on two separate afternoons within the same week.

The researcher conducted an initial session to introduce herself to each class. Administration procedures as outlined in the publication manuals were then followed.

The measures were delivered to the whole class simultaneously; the pupils were spaced apart and had a sheet to cover their responses. The researcher read out the instructions and all of the statements, as recommended, to control for differences in reading ability. Participants with low reading ages identified by the class teacher were seated together and received extra adult support. The papers were collected by the researcher and kept in a secure location.

For T2, the administration procedures were followed as identically as possible, over two afternoons in January 2011.

The teachers were asked to complete the T1 SDQs by October half term and the T2 SDQs by February half term, before the CG began the intervention. As two teachers job-shared the CG, it was ensured that the same teacher completed the SDQ for the same children at both time points.

3.13 Data Analysis

Analyses were conducted as follows to explore the effect of the intervention on the dependent variables (***Table 3.2***).

Dependent Variable	Measures
Emotional Distress	PI-ED Emotional Symptoms (SDQ)
Academic Self-Perceptions	MALS
Pupils' Strengths and Difficulties	SDQ: Emotional Symptoms Conduct Hyperactivity Peer Problems Prosocial Behaviour Total Difficulties

Table 3.2: Dependent variables and associated measures

Separate analyses of the SDQ subscales were conducted in addition to the Total Difficulties score. The Emotional Symptoms subscale complemented the analysis of 'emotional distress'; the Conduct and Hyperactivity scales informed an understanding of whether changes to internalising problems were accompanied by corresponding changes to externalizing problems; the Peer Problems and Prosocial scales indicated whether the peer learning aspect of FRIENDS impacted upon the children's behaviour. Although externalizing problems and prosocial behaviour are not identified explicitly within the research questions, their association with emotional distress was highlighted in the literature review. It was therefore considered that these additional analyses may illuminate an understanding of the processes operating in the intervention.

In order to answer research questions 1, 3 and 4, descriptive and inferential analyses of the data were performed. Chapter Four provides a full rationale for the choice of tests and procedures used in this study. These included tests of normality upon pre- and post-test scores (using z scores and Shapiro-Wilk tests; Conover, 1999); non-parametric analyses of initial differences

between groups (Mann Whitney U tests) and non-parametric analyses of each group's pre to post-test change (Wilcoxon tests). In addition, the researcher opted to perform a change score analysis (T2-T1) on the three measures (using non-parametric and, where justified, parametric procedures). This method is recommended widely for the use of quasi-experimental data because it controls, to some extent, for pre-existing differences between the groups that would not be adequately addressed by an analysis of (co)variance (Coolican, 2002; Maxwell and Delaney, 2004).

To answer research question 2 regarding preventive effects, an analysis of the number of children moving in and out of the 'at risk' category for emotional distress was conducted. A similar analysis of the children within the 'low academic self-concept' range was conducted for the MALS to complement question 3.

Finally, some post hoc correlations were conducted between particular sets of change scores to explore hypotheses emerging in the Discussion.

3.14 Ethical Considerations

This study has received full approval from the University of Nottingham's Ethics Committee. This section describes key ethical considerations as outlined by the British Psychological Society (2006), Health Professions Council (2008) and University of Nottingham (2009; ***Table 3.3***). Reference to specific items on the University of Nottingham's Ethical Checklist can be found in Appendix 3i.

Ethical Principles	Description [Adapted from BPS, 2006]	Methods employed to address issues
General Respect [BPS 1.1; HPC 1,3; UoN 3.2]	• Respect individual, cultural and role differences • Respect knowledge, insight, expertise of clients • Avoid unfair/prejudiced practice	• Participation of all children was valued and individual differences celebrated through the programme. • Knowledge, experience and expertise of school staff integrated into project.
Privacy and Confidentiality [BPS 1.2; HPC 2, 6; UoN 3.5, 3.6, 3.8, 4.1, 4.2, 4.3, 5.8]	• Obtain consent of clients for disclosure of confidential information • Record, process and store information securely • Ensure clients are aware of limitations of confidentiality • Restrict breaches of confidentiality to exceptional circumstances (eg. Concern about welfare) • Consult professional colleagues	• Consent obtained via parent letter regarding limitations and breaches of confidentiality • Children made aware of limitations verbally during administration instructions • Breaches restricted to cases of high scores on PI-ED at T2 in consultation with school/child and parents if necessary [Note limitations of PI-ED discussed in Chapter 5] • Personal data collected, stored and processed anonymously and securely. • Consultation sought from research supervisor

Informed Consent [BPS 1.3; HPC 7, 9, 10; UoN 3.10, 4.1.2]	• Ensure that clients understand the nature, purpose and anticipated consequences of participation • Obtain informed consent • Keep adequate records of consent • Avoid intentional deception of clients	• Written consent obtained from head teacher following presentation to staff about aims, purposes and requirements of research • Written consent obtained from parents following letter/attendance at meeting/discussion with family support worker • Pupils given basic information about study and offered opportunity to withdraw from participating in questionnaires
Self-Determination [BPS 1.4, HPC 1]	• Ensure clients' awareness of right to withdraw	• Staff and parents were informed in letters of right to withdraw children's data at any point
Limits of competence [BPS 2.3; HPC 6, 13; UoN 3.6]	• Practice within the boundaries of competence • Seek supervision when indicated	• Regular consultation with supervisor about issues causing concern (eg. large number of high anxiety scores).
Protection of Research Participants [BPS 3.3, HPC 1, 8; UoN	• Eliminate potential risks to psychological well-being • Inform participants when evidence is obtained of a psychological	• Programme focuses on improving well-being and therefore psychological harm is not implied • Worry box • Access to adult consultation and support

3.11]	problems that may endanger present or future well-being	• Booster groups to reinforce skills during Summer Term • Supervision provided to group leader • Right to withdraw
Debriefing [BPS 3.4; HPC 7; UoN 6.1]	• Debrief participants to inform about outcomes of research and arrange for further assistance as needed.	• Parents to be debriefed via a letter • Staff and pupils to be debriefed verbally (See Section 5:10) • Opportunities to clarify queries or misconceptions • Follow-up consultation for children causing concern if necessary

Table 3.3: A Description of Ethical Principles and Measures Taken to Address Them.

3.15 Pilot Study

The researcher conducted a pilot study with a Year 3 class in the participating school during the Summer Term 2010, (n= 23, 13 boys, 10 girls). The purpose was to pilot aspects of the main project, including obtaining parental consent, administering and scoring the measures and gaining a thorough knowledge and understanding of the FRIENDS programme.

The researcher worked alongside the class teacher (trained in FRIENDS) and the Learning Mentor who was to lead the main study. She supported 7 out of the 10 sessions and held weekly consultations with staff to gain their views on the programme. The data from this study was scored and examined for possible trends, but as the children were younger than the baseline standardisation age for the PI-ED and MALS and observed changes were small, it was not subjected to statistical analysis.

Results

Mean scores at T1 and T2 indicated the following trends:

- a very slight rise in self-reported levels of emotional distress (T1= 9.57, T2 = 10.95)
- stability in academic self-perceptions (T1 = 64.01; T2 = 64.26);
- a decrease in teacher-reported difficulties (T1= 8.83, T2 = 4.26)
- stability in teacher-rated prosocial scores (T1 = 9.04, T2 = 9.17).

The researcher acknowledged that drawing conclusions from these results was limited due to the young age group and the high likelihood of biased teacher reports due to lack of blinding. Anecdotal responses from both the pupils and teachers were extremely positive, however, and the researcher considered it justified repeating the process with an older age group who matched the standardisation profile of the measures more reliably. Appendix 3j details the key observations from the pilot study that influenced the main study implementation.

3.16 Reliability and Validity

As suggested in section 3.4.1, quasi-experiments are particularly vulnerable to a number of factors that threaten their reliability and validity (Cohen et al, 2007). 'Internal validity' refers to the plausibility of causal relationships demonstrated between treatment and outcome, while 'external validity' refers to the generalisability of results (Robson, 2002). 'Threats to validity' refer to the 'clouding conditions' that potentially interfere with these processes (Cohen et al, 2007). ***Table 3.4*** and ***Table 3.5*** illustrate the potential threats to internal and external validity in this study and, where possible, how the researcher has attempted to address them.

Threats to Internal Validity

Validity Threat	Description	How addressed
History	Things that have changed in the participants' environment other than those forming a direct part of the enquiry	Control group
Testing	Changes occurring as a result of practice/experience on pre-tests	N/A
Regression	Unusual or atypical scores at pre-test tend to become less unusual at post-test ('regression to the mean')	Attention to crossover analysis (Reichardt and Mark, 2001; see Chapter 5).
Mortality	Participants dropping out of the study	The data of 4 pupils who left the school before the post-test measures was excluded from the analysis.
Maturation	Growth, change or development in participants unrelated to treatment	Control group
Selection	Individual differences between groups prior to involvement	Change score analysis
Selection by maturation interaction	Predisposition of groups to grow apart (or together if initially different)	Attention to crossover analysis (Reichardt and Mark, 2001; see Chapter 5).
Diffusion of treatments	When the control/comparison groups inadvertently receives aspects of the treatment	Quasi-experimental design; details of FRIENDS withheld from the CG class and teacher
Compensatory equalization of treatments	If one group receives 'special' treatment, there will be organization and other pressures for the control to receive it Robson (2002, p105-6), after Cook and Campbell (1979, pp51-5)	Control group to receive the treatment in Summer Term so threat minimized.

Table 3.4: Threats to Internal Validity and how the researcher has addressed them.

Threats to External Validity

Validity Threat	Description	How addressed
Lack of representativeness of available and target populations	Sample population may not be representative of the wider population to which the experimenter seeks to generalise the findings	N/A
Hawthorne Effect	The psychological effects that arise out of participating in an intervention	N/A *Blinding not possible in this study*
Inadequate operationalising of the dependent variable	Whether the methods employed accurately reflect the constructs under scrutiny	Measures focus on the dependent variables in an educational context
Sensitization / reactivity to experimental conditions	See 'Testing' in Table 3.4.	
Invalidity / unreliability of instruments	Data is yielded in which confidence cannot be placed	Published measures used with a degree of established reliability and validity (see Chapter 5 for a discussion of limitations of the PI-ED)
Ecological Validity	The extent to which behaviour in one context can be generalised to another After Cohen et al (2007, *p156-7*)	Quasi-experimental design preserving intact groups

Table 3.5: Threats to External Validity and how the researcher has addressed them.

3.17 Strengths and Limitations

This study demonstrates strengths in its application of contemporary models of EP practice; by undertaking an evaluation of a programme that is reportedly popular, but that remains under-investigated in controlled studies in the UK. The researcher has a sound knowledge of the programme components and has been in a strong position to reinforce the underlying psychological principles. This has also enabled her to contribute to the evaluation of treatment integrity and to the promotion of skills, as well as considering the possible underlying mechanisms that might be implicated in effecting change. As well as evaluating the FRIENDS programme, the study also provides a working model for how EPs might work collaboratively with schools to promote mental health initiatives.

However, the quasi-experimental design entails a number of threats to internal and external validity. As will be illustrated in Chapter Four, the lack of random allocation entails significant limitations for the data analysis and subsequent ability to draw causal inferences in relation to the research questions. The related issue of selection bias in a small sample and the Hawthorne effect due to lack of blinding are also salient in this context, while the influence of experimenter bias has been raised in Section 3.9. Furthermore, the timing of the intervention, at the start of an academic year, may have influenced self-reports of anxiety, in addition to the limitations cited in section 3.11. As with all real world studies, there is also a strong likelihood of interaction effects between these various sources of bias, (Robson, 2002; Cohen et al, 2007).

Finally, the pupil outcomes evidenced in the quantitative data reported in this study represent a 'snapshot' of the effects of this project implementation. As the intervention proceeded, the researcher became increasingly aware, through anecdotal evidence, of its effects upon the children's behaviour, the teacher's perception of the pupils and the social relationships within the classroom. It was informally hypothesised that these factors were all likely

'mechanisms' in producing effects upon the dependent variables. In this respect, the potential for this investigation 'outgrew' its fixed design, although the scope of the study has not permitted a detailed exploration of these other factors. For these reasons and given the limitations of questionnaires, the researcher fully acknowledges the value of adopting mixed methods paradigms in future evaluation research, (Mertens, 2010).

4 CHAPTER FOUR: RESULTS

Introduction

In this chapter, a descriptive and inferential analysis of the pre and post-test data will be presented in order to address the research questions stated at the end of Chapter Two. These involved considering the effects of the intervention upon: the reduction and prevention of emotional distress (measured by self and teacher report); pupil academic self-perceptions (measured by self-report), and pupil behaviour (measured by teacher report). The data is organised according to the three measures used to evaluate these dependent variables: the Paediatric Index of Emotional Distress, the Myself-As-Learner Scale and the Strengths and Difficulties Questionnaire, which includes both a 'Total Difficulties' score (combining Emotional Symptoms, Conduct, Hyperactivity and Peer Problems) and individual subscale (including Prosocial) scores. The purpose of the analysis is to consider whether any of the experimental hypotheses outlined in section 2.6.1 may be supported and the null hypotheses rejected.

Table 4.1 illustrates the expected direction of change according to the experimental hypotheses:

Name of instrument	Sub-scale scores		Anticipated direction of change
	Pupil	*Teacher*	
PI-ED	Emotional distress	-	Decrease in scores
MALS	Academic self-perceptions	-	Increase in scores
SDQ	-	Emotional Symptoms	Decrease in scores
		Conduct	Decrease in scores
		Hyperactivity	Decrease in scores
		Peer problems	Decrease in scores
		Total Difficulties	Decrease in scores
		Prosocial	Increase in scores

Table 4.1: Measures, dependent variables and predicted directions of change.

4.1 Descriptive Analysis

This section presents descriptive statistics for each of the dependent variables. ***Table 4.2*** contains descriptive data for emotional distress, measured by the PI-ED, ***Table 4.3*** illustrates data for academic self-perceptions, measured by the MALS, and ***Table 4.4*** shows the analysis for behaviour measured by the teacher version of the SDQ. Because the inferential analysis will include both parametric and non-parametric analyses, (see section 4.3 for an explanation), the median and range for each measure will be included, in addition to means and standard deviations. Further pertinent output from PASW including confidence intervals for parametric tests can be found in Appendix 4. This section also includes graphic representation of the changes in mean scores between pre- and post-tests for each group to illustrate some general tendencies in the direction of the data. It should be noted, however, that care must be taken with using the mean as a measure of central tendency with the ordinal level data obtained in this study, as the intervals between participant responses are not necessarily equal (see section 4.3 for further explanation and discussion of this issue). However, many researchers in psychology treat Likert-scale data as interval level and thus justify the use of the mean in data analysis (www.researchmethodsinpsychology.com). The present researcher proceeds with caution, highlighting obvious discrepancies between mean and median values and acknowledging the effects of extreme scores on the calculation of the mean. (Appendices 4a, 4b and 4c show box and whisker plots to illustrate the spread of data and extreme scores for the PI-ED, MALS and Total Difficulties scores to assist with answering the main research questions).

Emotional Distress (PI-ED)

Time	Group	N	Mean	St. Dev.	Median	Range
Pre-test	FRIENDS	18	17.94	8.98	17.50	30.00 (4.00-34.00)
	Control	20	14.75	8.20	12.00	28.00 (4.00-32.00)
Post-test	FRIENDS	18	13.28	5.98	12.50	20.00 (4.00-24.00)
	Control	20	15.35	8.86	17.50	29.00 (1.00-30.00)

Table 4.2: Distribution and spread of data for the PI-ED

Observations

- The IG's pre-test mean and median scores are higher than the CG's but at post-test the IG's scores are lower than the CG's.

- Standard deviation values indicate that the spread of scores is relatively similar at pre-test, but by post-test, the IG's scores are clustered more closely around the mean.

- The range of scores is similar between both groups at pre-test; at post-test, the CG's range remains relatively stable but the IG's maximum score reduces considerably, indicating less extreme high values for this group, post-intervention,

- There are larger discrepancies between the CG's mean and median values than the IG's. The Box Plot in Appendix 4a illustrates how the spread of scores became greater for the CG.

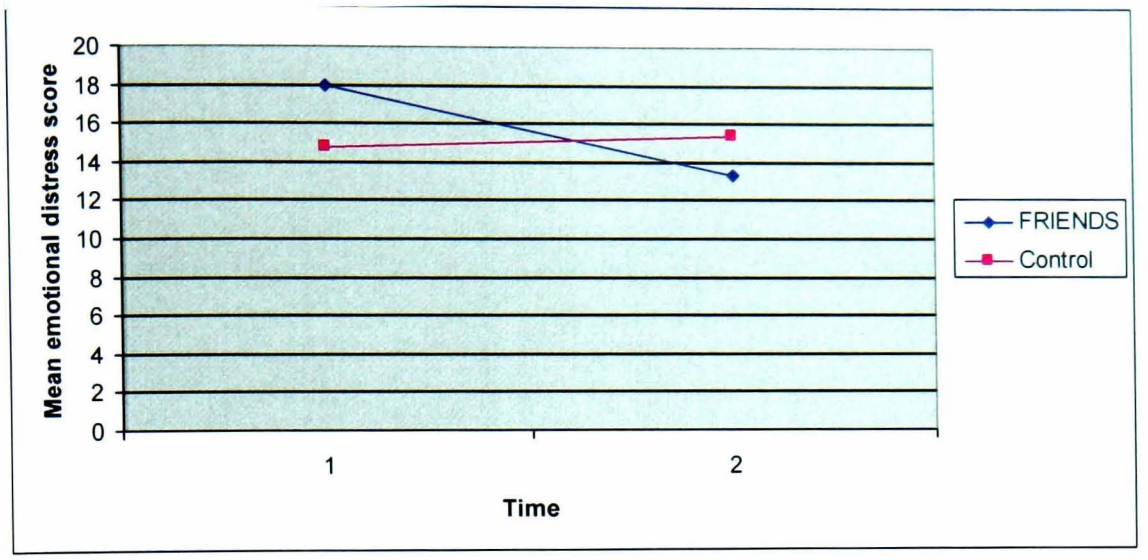


Figure 4.1: Mean emotional distress scores for IG and CG at pre and post-test

42 is the maximum score on this test.

Figure 4.1 illustrates that emotional distress scores for the IG decreased between pre and post-test, while the CG's scores showed a slight increase over the same period. This demonstrates a crossover interaction where the IG's mean score starts higher than the CG's but ends lower (Reichardt and Mark, 2001) and reflects the assumptions of the experimental hypothesis that IG emotional distress scores would decrease in comparison to CG scores. These scores will be analysed further in Section 4.5.

Academic Self-Perceptions (MALS)

Time	Group	N	Mean	St. Dev.	Median	Range
Pre-test	FRIENDS	18	66.61	18.58	68.00	72.00 (28.00-100.00)
	Control	20	64.90	14.56	65.00	61.00 (32.00-93.00)
Post-test	FRIENDS	18	71.94	10.99	68.50	47.00 (53.00-100.00)
	Control	20	64.75	13.94	65.00	45.00 (41.00-86.00)

Table 4.3: Distribution and spread of data for the MALS

Observations

- Mean scores at pre-test are similar between the two groups; at post-test the IG's mean score shows a slight rise, while the CG's score remains stable. Scrutiny of the box and whisker plots in Appendix 4b indicates that the IG's mean post-test score may have been elevated by the extreme score of participant 12.

- Standard deviation scores indicate a wider spread for the IG at pre-test; at post-test the IG's scores are clustered more closely around the mean while the CG's spread remains relatively stable.

- Both groups show a diminished range of scores at post-test with the minimum score rising in both groups. The IG illustrates at least one extreme maximum score at both pre and post-test.

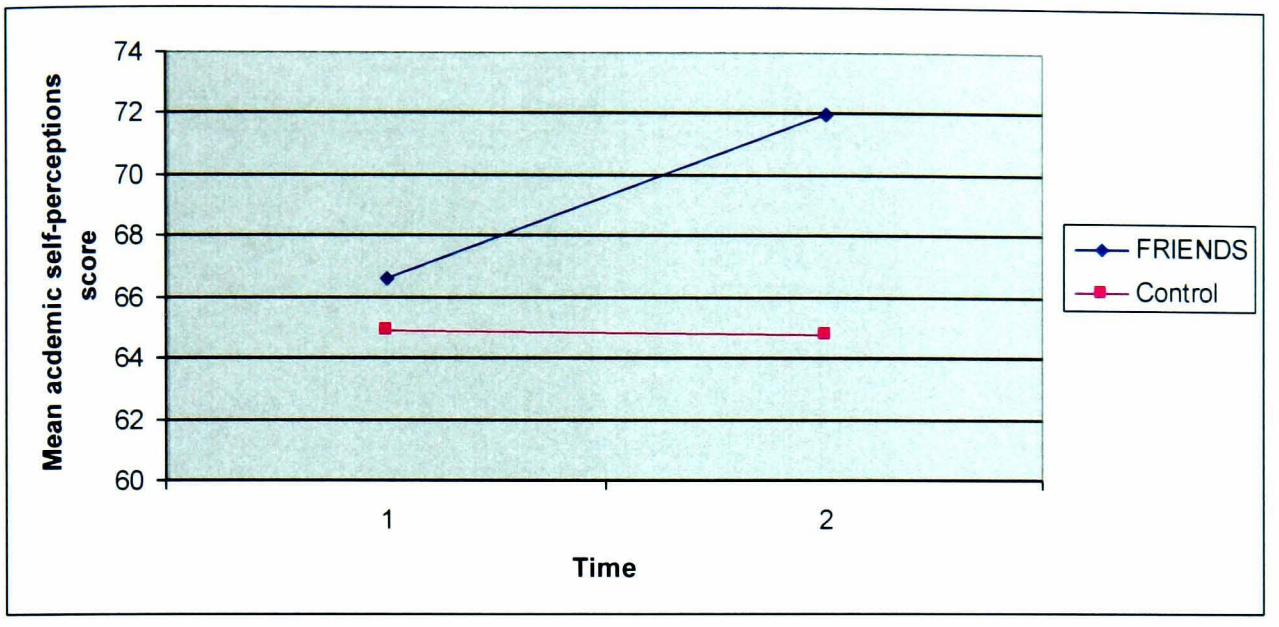


Figure 4.2: Mean academic self-perception scores for the IG and CG at pre and post-test

100 is the maximum score on this test.

Figure 4.2 illustrates that IG scores showed an increase in academic self-perceptions while CG scores remain stable over the same period. This relationship is in line with the research hypothesis that IG academic self-perceptions will increase as a result of the intervention in comparison to the CG. However, the IG's post-test mean may have been affected by an extreme score as highlighted previously. These scores will be analysed further in Section 4.5.

Strengths and Difficulties (SDQ)

Time	Subscale	Group	N	Mean	St. Dev.	Median	Range
Pre-test	**Total Difficulties**	**FRIENDS**	**18**	**10.67**	**6.53**	**10.00**	**20.00 (1.00-21.00)**
		Control	**20**	**12.50**	**8.85**	**12.00**	**30.00 (0.00-30.00)**
	Emotional Symptoms	FRIENDS	18	2.61	2.06	2.00	7.00 (0.00-7.00)
		Control	20	2.90	2.45	3.00	8.00 (0.00-8.00)
	Conduct	FRIENDS	18	1.44	2.23	0.50	7.00 (0.00-7.00)
		Control	20	2.55	2.78	2.00	9.00 (0.00-9.00)
	Hyperactivity	FRIENDS	18	4.78	2.98	4.00	10.00 (0.00-10.0)
		Control	20	5.40	3.87	6.00	10.00 (0.00-10)
	Peer Problems	FRIENDS	18	1.83	1.95	1.00	5.00 (0.00-5.00)
		Control	20	1.65	1.27	2.00	5.00 (0.00-5.00)
	Prosocial	FRIENDS	18	6.33	2.97	7.00	9.00 (1.00-10.00)
		Control	20	6.15	3.03	5.50	9.00 (1.00-10.00)
Post-test	**Total Difficulties**	**FRIENDS**	**18**	**6.00**	**3.25**	**6.00**	**12.00 (0.00-12.00)**
		Control	**20**	**10.10**	**7.67**	**8.50**	**23.00 (0.00-23.00)**
	Emotional Symptoms	FRIENDS	18	1.06	1.21	1.00	4.00 (0.00-4.00)
		Control	20	1.75	1.45	1.00	5.00 (0.00-5.00)
	Conduct	FRIENDS	18	0.89	1.02	1.00	3.00 (0.00-3.00)
		Control	20	1.55	2.19	1.00	7.00 (0.00-7.00)
	Hyperactivity	FRIENDS	18	2.89	1.94	3.00	6.00 (0.00-6.00)
		Control	20	5.25	3.73	5.00	10.00 (0.00-10)
	Peer Problems	FRIENDS	18	1.17	1.42	1.00	4.00 (0.00-4.00)
		Control	20	1.55	1.70	1.00	5.00 (0.00-5.00)
	Prosocial	FRIENDS	18	8.06	2.29	9.00	7.00 (3.00-10.00)
		Control	20	7.15	2.37	7.50	7.00 (3.00-10.00)

Table 4.4: Distribution and spread of data for the SDQ subscales

Observations

- Mean scores for Total Difficulties show a decrease for both groups, with the IG showing the larger change. Standard deviation values indicate that the IG shows a narrower spread of data at pre-test and this diminishes further at post-test. Both groups show a reduction in the range of scores with maximum scores being considerably lower at post-test. The box plots in Appendix 4c illustrate these observations.

- For Emotional Symptoms, both groups show a decrease in mean, median and range scores from pre-test to post-test with no discernible differences between the two.

- Similarly, both groups' mean and range scores for Conduct show a decrease at post-test with the IG having the lower score at both time points. The range of conduct scores for the CG is greater with a higher maximum score at both pre and post-test.

- For Hyperactivity, the groups have similar mean scores and identical ranges (0-10) at pre-test. At post-test, the IG's mean score has reduced while the CG's score has remained relatively stable. The IG's range of scores has similarly decreased, with a lower maximum score, while the CG's has stayed the same.

- Mean, standard deviation, median and range scores for Peer Problems indicate relative stability both within and between groups at both time points.

- Mean and median Prosocial scores show an increase for both groups from pre-test to post-test. Standard deviation values indicate relative stability between and within groups for this domain.

The changes in mean scores for the SDQ subscales will now be illustrated graphically. The scale on each graph has been standardised to permit a

clearer comparison between the variables (10 is the maximum score on the individual subscales and 40 is the maximum score for Total Difficulties). All of these scores will be analysed further in Section 4.5.

Total Difficulties

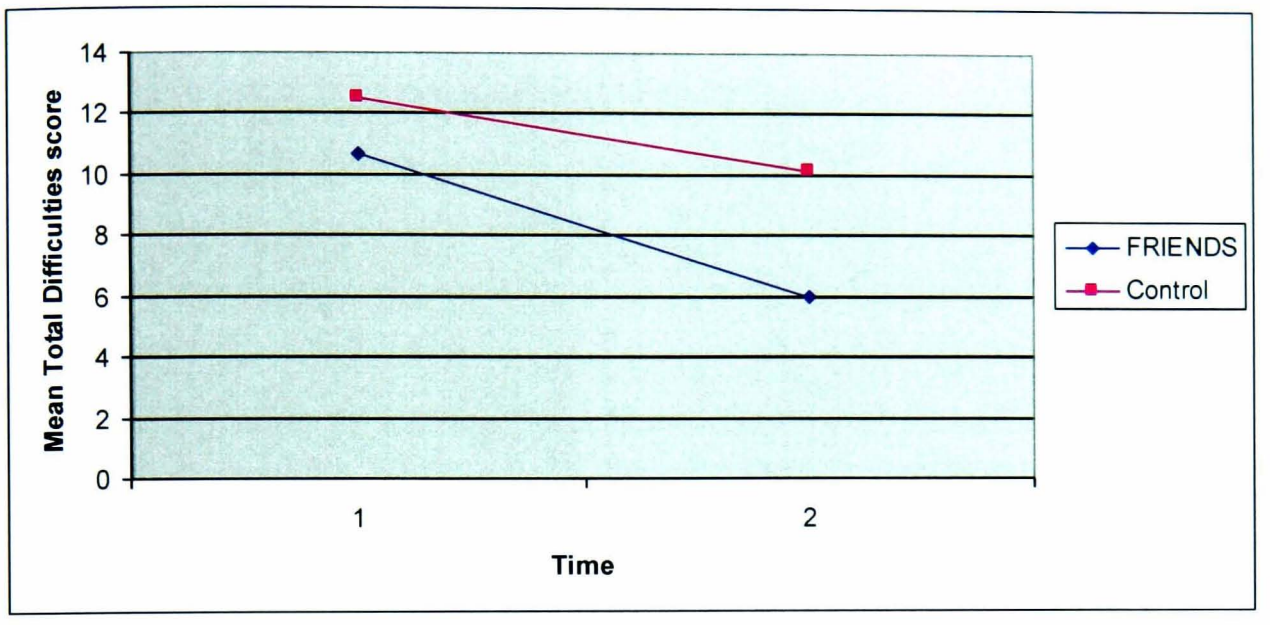


Figure 4.3: Mean teacher-rated Total Difficulties scores between the IG and CG at pre and post-test.

The maximum score on this test is 40.

Figure 4.3 illustrates that teacher-rated Total Difficulties for both the IG and CG decreased between pre and post-test. The mean score for the IG was lower at pre-test and appeared to fall slightly more sharply than for the CG, in line with the hypothesis of reduced overall difficulties as a result of the intervention. However, as a reduction in the mean scores for both groups is observed, it will be necessary to conduct inferential analyses to detect any significant differences in change between the two (see section 4.5).

Emotional Symptoms

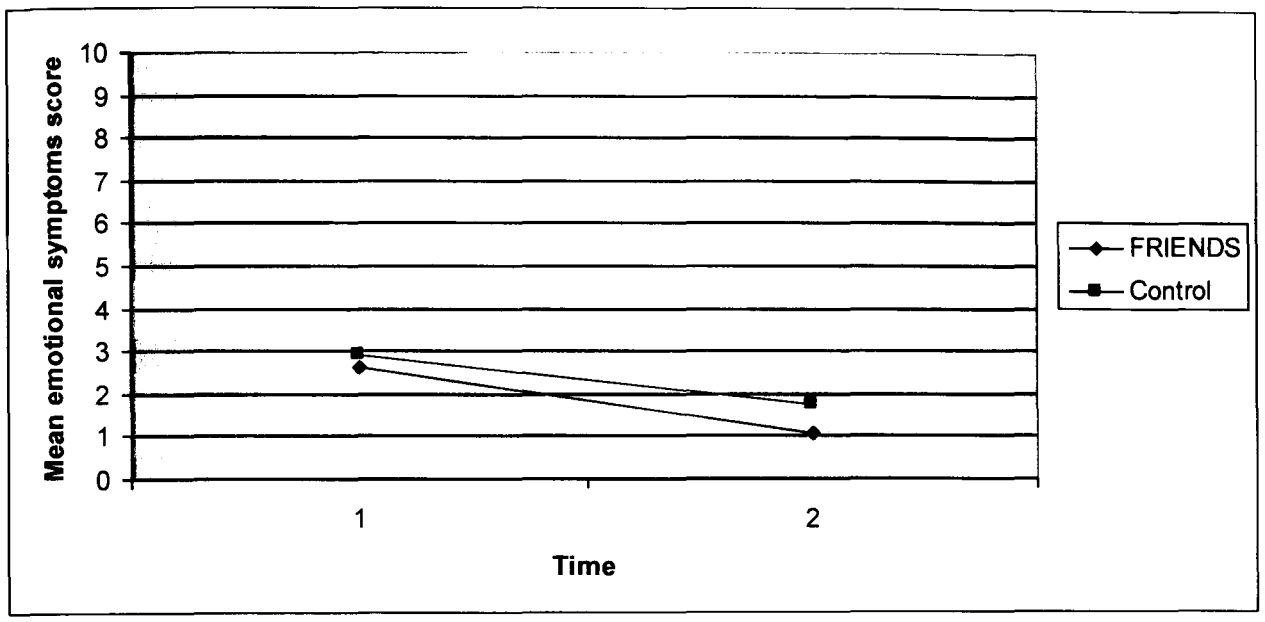


Figure 4.4: Mean teacher-rated emotional symptoms scores on the SDQ for IG and CG at pre and post-test.

Figure 4.4 illustrates that mean teacher rated scores of emotional symptoms for both the IG and CG were relatively low at the start and decreased between pre and post-test. Although the direction of change is correctly predicted by the hypothesis that emotional symptoms would decrease, the IG does not appear to have shown a greater decrease in scores than the CG as the gradient of both lines is similar.

Conduct

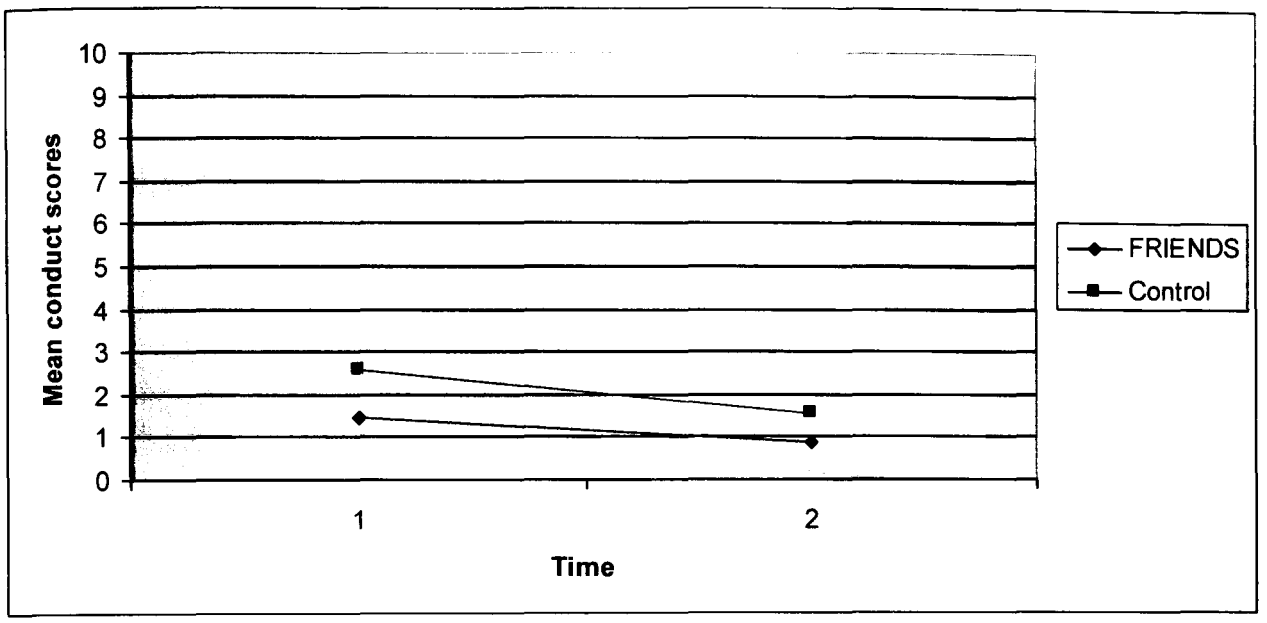


Figure 4.5: Mean teacher-rated scores for behaviour problems on the SDQ for IG and CG at pre and post-test.

Figure 4.5 illustrates that mean teacher-rated scores for conduct problems were in the low range at the start and decreased for both the IG and CG between pre and post-test. Although the direction of change is again predicted by the hypothesis that pupil conduct difficulties would decrease, the IG does not appear to have demonstrated a greater decrease than the CG as the lines have similar gradients.

Hyperactivity

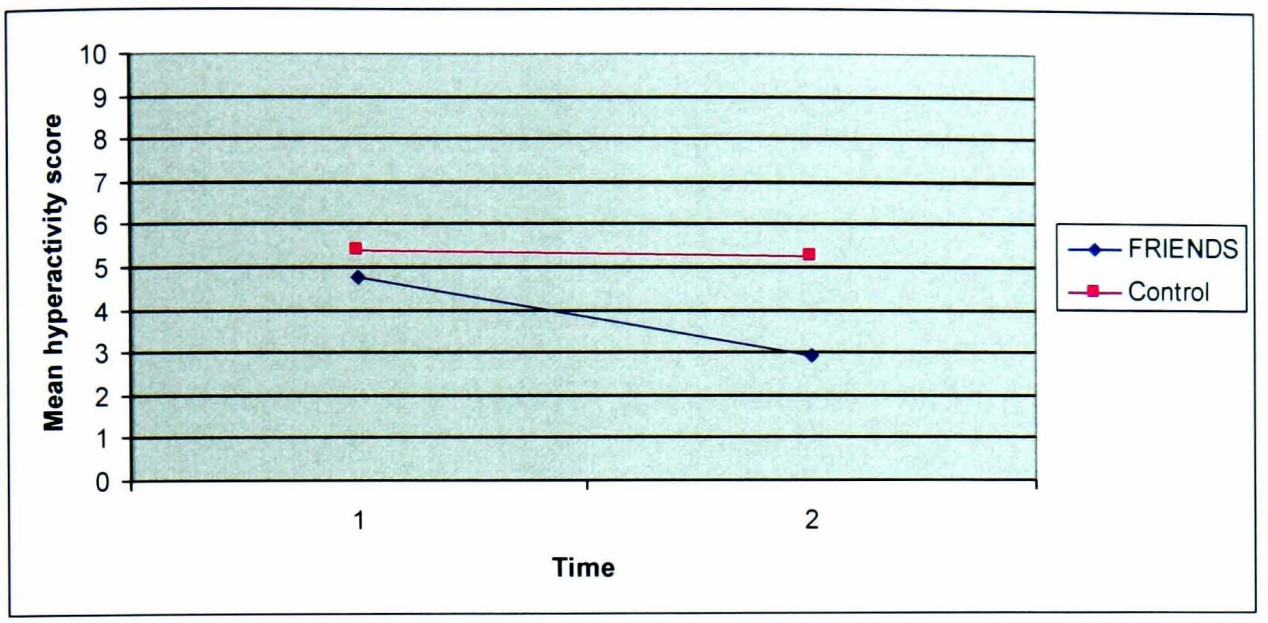


Figure 4.6: Mean teacher-rated hyperactivity scores on the SDQ for IG and CG at pre and post-test.

Figure 4.6 illustrates that mean teacher ratings of hyperactivity were within the mid-range at pre-test. Both groups showed a decrease in teacher-rated hyperactivity scores, with the IG showing an apparently greater change than the CG, as indicated by the gradient of the blue line.

Peer Problems

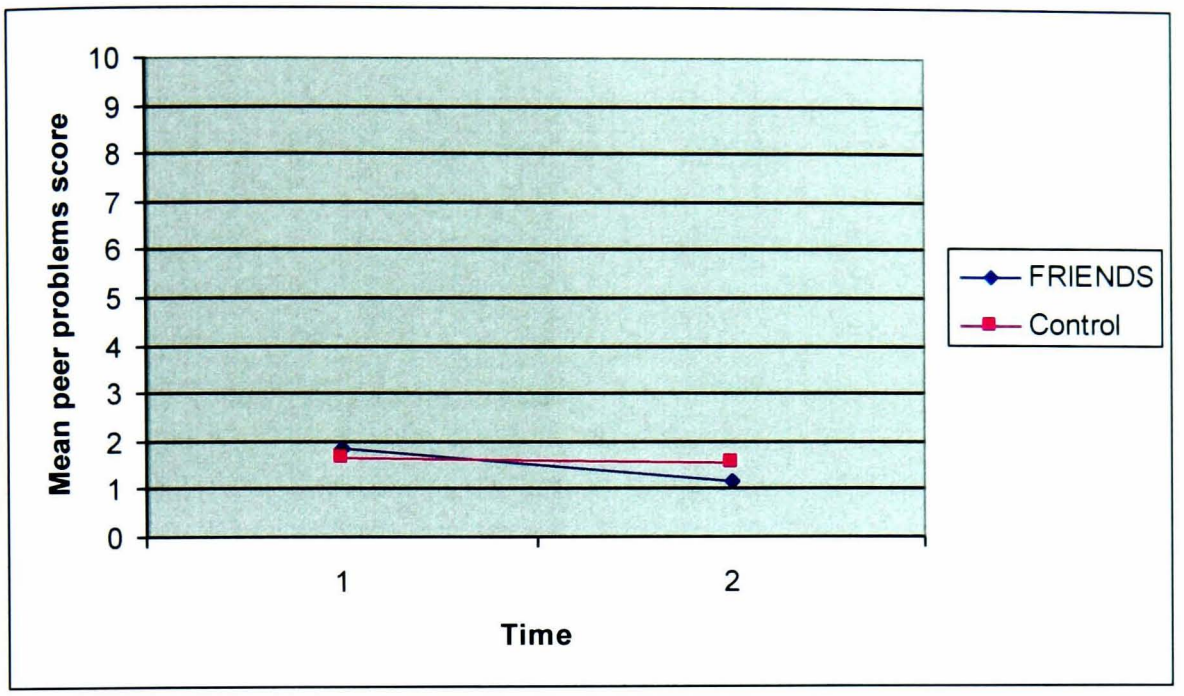


Figure 4.7: Mean teacher-rated peer problem scores for IG and CG at pre and post-test.

Figure 4.7 illustrates that mean teacher ratings of peer problems were within the low range at the start. Both the IG and CG showed decreases in scores with the IG showing a slightly larger change than the CG resulting in a small crossover effect. This supports the hypothesis that the IG scores would show a greater decrease in teacher-rated peer problems, although the discrepancy in gradients between the groups is minimal.

Prosocial

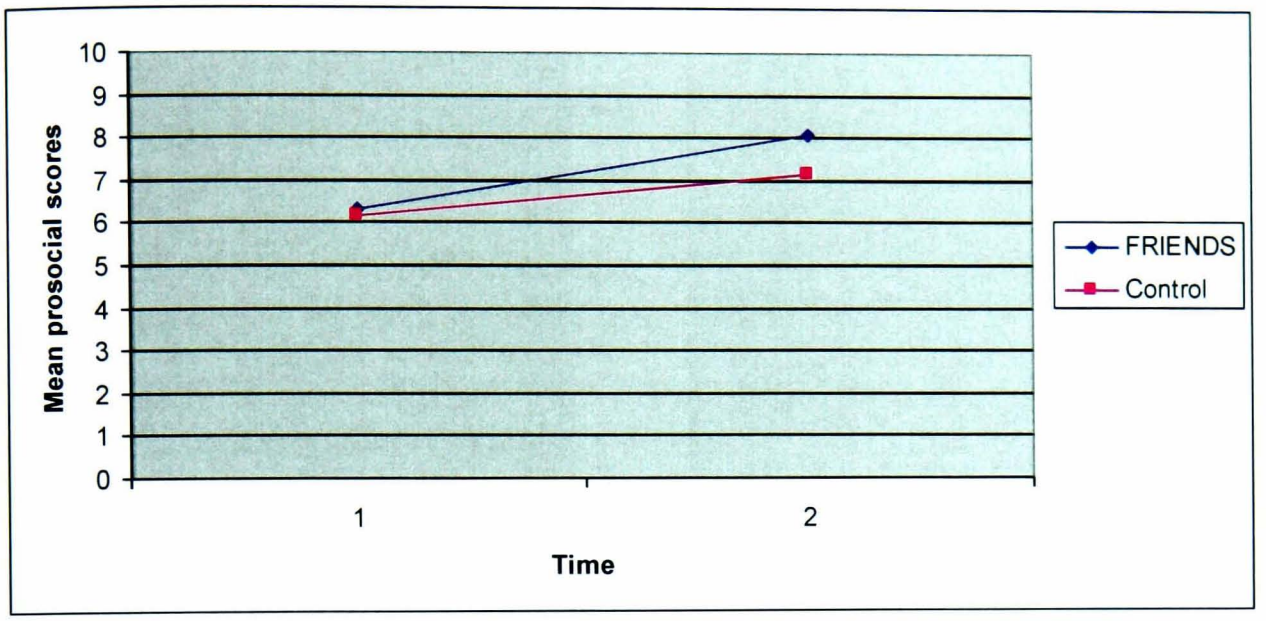


Figure 4.8: Mean teacher-rated Prosocial scores for IG and CG at pre and post-test.

Figure 4.8 illustrates that both groups' mean scores were similar at pre-test, within the mid to high range, and both increased at post-test. The gradient of the blue line is slightly steeper, indicating a greater change for the IG, which supports the hypothesis regarding the development of peer relations during the intervention. However, as both scores showed an increase, it will be necessary to conduct inferential analyses to detect any significant differences between the groups' improvement (see Section 4.5).

4.1.1 Change Scores

For reasons explained later in section 4.3, an analysis of the change scores for each group was conducted in order to compare the amount of change between the groups. The change score is calculated by subtracting the T2 score from the T1 score and this then becomes the dependent variable on which the analysis is conducted (Gliner et al, 2003).

Results for the change scores were as follows:

Emotional Distress (PI-ED)

Group	N	Mean change	St. Dev.	Median	Range
FRIENDS	18	-4.67	8.09	-3.00	34.00 (-25.00-9.00)
Control	20	.75	6.89	1.00	25.00 (-10.00-15.00)

Table 4.5: Distribution and spread of data for the emotional distress change scores

The IG's mean and median change scores constitute negative values indicating a reduction in emotional distress while the IG's are positive indicating a slight rise.

Academic Self-Perceptions (MALS)

Group	N	Mean change	St. Dev.	Median	Range
FRIENDS	18	5.83	13.91	4.00	57.00 (-21.00-36.00)
Control	20	-.15	10.29	-1.00	42.00 (-25.00-17.00)

Table 4.6: Distribution and spread of data for the academic self-perceptions change scores

The IG's mean and median change scores are positive values indicating a rise in academic self-perceptions while the CG's negative values indicate a slight decrease.

Behaviour (SDQ)

Total Difficulties

Group	N	Mean change	St. Dev.	Median	Range
FRIENDS	18	-4.67	4.31	-4.50	16 (-15.00-1.00)
Control	20	-2.40	3.97	-3.50	15 (-9.00—6.00)

Table 4.7: Distribution and spread of data for the Total Difficulties change scores

Mean and median change scores indicate a reduction in Total Difficulties for both groups.

Emotional Symptoms

Group	N	Mean change	St. Dev.	Median	Range
FRIENDS	18	-1.56	1.34	-1.00	5.00 (-5.00-0.00)
Control	20	-1.15	2.60	-.50	10.00 (-7.00-3.00)

Table 4.8: Distribution and spread of data for the Emotional Symptoms change scores

The mean and median change scores for both groups constitute negative values indicating a slight reduction in emotional symptoms for both groups.

Conduct

Group	N	Mean change	St. Dev.	Median	Range
FRIENDS	18	-.56	1.62	0.00	5.00 (-4.00-1.00)
Control	20	-1.00	1.45	0.00	4.00 (-4.00-0.00)

Table 4.9: Distribution and spread of data for the Conduct change scores

The mean scores indicate a very slight negative trend in behaviour problems for both groups, although the medians suggest no change.

Hyperactivity

Group	N	Mean change	St. Dev.	Median	Range
FRIENDS	18	-1.89	2.63	-1.50	9.00 (-7.00-2.00)
Control	20	-.15	1.53	0.00	6.00 (-3.00-3.00)

Table 4.10: Distribution and spread of data for the Hyperactivity change scores

The mean and median change scores for the IG indicate a reduction in hyperactivity in comparison to the CG, which remains relatively stable.

Peer Problems

Group	N	Mean change	St. Dev.	Median	Range
FRIENDS	18	-.67	1.28	-.50	5.00 (-3.00-2.00)
Control	20	-.10	1.12	.00	4.00 (-2.00-2.00)

Table 4.11: Distribution and spread of data for the Peer Problems change scores

Mean and median change scores indicate a slightly more negative trend for the IG than the CG.

Prosocial

Group	N	Mean change	St. Dev.	Median	Range
FRIENDS	18	1.72	2.24	1.00	9 (-1.00-8.00)
Control	20	1.00	1.49	1.00	5.00 (-1.00-4.00)

Table 4.12: Distribution and spread of data for the Prosocial change scores

Mean and median scores indicate a slight rise in prosocial behaviour for both groups.

Section 4.2 onwards comprises an inferential analysis based on the above data in order to test for any significant differences in change between the IG and CG and ultimately to answer the research questions posed. The final part of this section will focus on the question regarding preventive effects and the evidence for these in the present study.

4.1.2 Preventive Effects

In order to answer the research question regarding preventive effects, the researcher has followed the protocol used in other FRIENDS studies (Gallegos, 2008; Stallard et al, 2008), which involves reporting the number and percentage of children moving in and out of the elevated range of scores on the continuum of emotional distress. A calculation of the number of children moving in and out of the range of 'low academic self-concept' has also been conducted to inform research question 3.

Because of the limited sample size involved, it was not deemed appropriate to conduct statistical analyses on the results. (Mertens (2010), recommends a minimum of about fifteen participants per variable for statistical tests).

Emotional Distress

Sensitivity and specificity co-efficients for the PI-ED revealed cut-off values of 10 for boys and 11 for girls (O'Connor et al, 2010). Scores above these values are considered to be within the range of 'clinical concern' for emotional distress. When the researcher applied these values to the initial T1 data, however, a very large proportion of both classes scored above these suggested cut-off points. A subsequent consultation with one of the test's authors revealed that further data was being collected and this was a common finding, and it was therefore likely that the score would be modified in the future to at least 15, (Personal communication with E. Ferguson, 01.04.11).

Evidence suggests that the skewed distributions and violations of assumptions of normality in general population samples may distort the calculation of 'clinical' cut off points (Martinovich et al, 1996). This has led researchers such as Connell and colleagues (2007) to conclude that cut-off scores should be "used thoughtfully and adjusted to fit context and purpose" (*p69*). The implications of these ambiguities will be referred to in Chapter Five. For the purpose of the present analysis, the suggested score of 15 will be applied as a threshold from which to assess the movement of children within

'higher' and 'lower' ranges of scores, although the intention is not to represent 'clinical' categories here.

Group	Total / % 'at risk' at T1	Total / % 'at risk' at T2	Total / % staying in 'at risk' category at T2	Total / % moving out of 'at risk' category at T2	Total / % moving into 'at risk' category at T2
FRIENDS	10 55.55%	6 33.3%	5 27.5%	5 27.5%	1 5.5%
Control	7 35%	11 55%	7 35%	0 0%	4 20%

Table 4.13: Number and percentage of children moving in and out of the higher range of scores in the IG and CG.

Summary

Table 4.13 illustrates that over half of the IG started in the 'at risk' category at pre-test but this decreased to around a third at post-test. Half of the children 'at risk' at T1 moved out of this category at T2 and one child moved in. (Of the five children who stayed in the higher category, four had reduced scores at T2).

In contrast, the CG showed a different trend, with around a third of the class being 'at risk' at pre-test, moving to over a half at post-test. All of the children who were 'at risk' at T1 stayed in this category at T2 (and of these seven children, five had increased scores at post-test). Furthermore, four children moved into the 'at risk' category at T2.

These results provide tentative evidence for a prevention effect as a result of the intervention, which will be discussed in Chapter Five.

Academic Self-Perceptions

A score of below 60 on the MALS represents the range of low academic self-perceptions (Burden, 1998).

Group	Total / % in low self-perception range at T1	Total / % in low self-perception range at T2	Total / % staying in low self-perception range at T2	Total / % moving out of low self-perception range at T2	Total / % moving into low self-perception range at T2
FRIENDS	6 33%	1 5.5%	1 5.5%	5 27.8%	0 0%
Control	5 25%	7 35%	4 20%	1 5%	3 15%

Table 4.14: Number and percentage of children moving in and out of the lower range of scores in the IG and CG.

Summary

Table 4.14 illustrates that around a third of the IG started in the 'low academic self-perception' range, reducing to 5.5% at post-test. 5 children moved out of this range and no children moved in.

Around a quarter of the CG started in the 'low academic self-perception group', rising to over a third at post-test. Of the original five children who were in the low range at T1, four remained at this status at T2 with just one child moving out and a further three moving in.

These results provide tentative evidence for a beneficial effect of the intervention for those children with low academic self-perceptions and a possible prevention effect for other children moving into this range.

4.2 Inferential Analysis

Rationale

In contrast to descriptive analyses that merely report what has been found, inferential statistics strive to make inferences and predictions based on the data gathered (Cohen et al, 2007). This section provides a rationale for the researcher's choice of inferential analyses, alongside descriptive statistics, which were used to answer the research questions posed in the present study. Table 4.15 shows how the data was applied to answering the questions.

Research Question	Data used	Type of analysis
1. Does a class of Key Stage 2 children participating in a universal FRIENDS intervention report a significant reduction in emotional distress (ED) in comparison to a non-intervention control group?	- PI-ED self-report - SDQ Emotional Symptoms teacher report	- Inferential analysis of mean, median and change scores
2. Does a class of Key Stage 2 children participating in a universal FRIENDS intervention experience a preventive effect for ED in comparison to a non-intervention control?	- PI-ED self-report	- Descriptive analysis of ED 'prevention' scores
3. Does a class of Key Stage 2 children participating in a universal FRIENDS intervention display significantly more positive academic self-perceptions (ASP) than those in a non-intervention control group?	- MALS self-report	- Inferential analysis of mean, median and change scores - Descriptive analysis of ASP 'prevention scores'
4.Does participation in a universal FRIENDS programme result in a significant improvement in teacher-rated pupil behaviour in comparison to a non-intervention control?	- SDQ teacher-report	- Inferential analysis of 'total difficulties' and individual subscale scores

Table 4.15: Table to show data and type of analysis used to answer each research question.

4.3 Issues Influencing the Choice of Analysis

To answer the research questions, the researcher conducted a quasi-experimental, non-equivalent groups design on a sample of 38 children. This involved using three measures, all based on Likert-type rating scales, which produced ordinal level data (see Cohen et al, 2007, for a review of levels of data, *p502*).

In deciding on the method of inferential analysis to undertake, it was necessary to consider the nature of the design and the level of data obtained. In studying these aspects, the researcher identified a number of difficulties associated with the arrangements in the present study.

4.3.1 Problems with Interpreting Ordinal Level Data

As highlighted previously, data obtained from questionnaires and surveys may be classified as ordinal level. This means that cases may be ranked according to a numerical order on the variable from the smallest to the largest, but there is an absence of calibrated or equal intervals between the items; as opposed to 'interval' or 'ratio' level data (Howitt and Cramer, 2011; Cohen et al, 2007). As described in section 4.1, this suggests that the median and range should be reported as measures of central tendency for ordinal level data as opposed to means and standard deviations (Gay, Mills and Airasian, 2009). Another implication is that, because ordinal data is considered to be 'non-parametric', in other words, no assumptions can be made about the characteristics of the underlying population, it is inappropriate to apply parametric procedures to this type of data (Jamieson, 2004). This has led certain authors to suggest that ordinal data should be rescaled to interval data to satisfy assumptions of normality (Harwell and Gatti, 2001).

However, this topic is a source of considerable debate within the research community. As Norman (2010) argues, if parametric tests on Likert-scale data were disallowed, then a large proportion of research on educational, health status and quality of life assessment would be dismissed. Norman

(2010) proceeds to demonstrate how the versatility and robustness of parametric tests address the issues of skewness, non-normality and nonlinear relations with respect to ordinality. He concludes that, consistent with a range of empirical data from the past 80 years, parametric tests may be used with Likert data, small sample sizes and non-normal distributions with no fear of "coming to the wrong conclusion." This view was echoed by Velleman and Wilkinson (1993), who challenged the validity and usefulness of Stevens's (1946, *cited by ibid.*) taxonomy in selecting statistical methods. They emphasised that good data analysis involves searching for interesting patterns and unanticipated relationships and that approaching analysis from an *a priori* scale type that excludes certain statistical procedures may limit the kinds of hypotheses and discoveries to be made.

The implications of these observations have informed the researcher's choice of analyses, which will be discussed later.

4.3.2 Problems Associated with Non-Equivalent Groups Designs

Authors highlight that an essential feature of the pre-test/post-test comparison group design is the random allocation of participants (Gliner et al, 2003; Trochim, 2006). For reasons explained in section 3.6.1, this protocol was not followed in the present study and is indeed, common practice in many educational evaluations (Cohen et al, 2007, *p282*). However, a problematic consequence of this design is that it has significant implications for the type of data analysis that can be used. Authors generally concur that applying traditional parametric measures, such as the Analysis of Covariance (ANCOVA), to compare pre and post-test scores with non-equivalent groups is essentially flawed (Gliner et al, 2003; Trochim, 2006). This is because the non-randomization incurs certain types of bias: for example, the covariate adjustment in ANCOVA can introduce false relationships between group assignment and outcome (Fitzmaurice, Laird and Ware, 2004). This may lead to a Type II error (accepting the null hypothesis when there is, in fact a difference between the groups), as the covariate may explain away meaningful differences (Smolkowski, 2010). Conversely, Reichardt and Mark

(2001) argue that the selection bias inherent in quasi-experimental designs may lead to a Type I error, resulting in apparent post-treatment difference when there is in fact, no treatment effect. Both of these issues are controlled more successfully by the randomisation of participants (Cohen et al, 2007, *p155*), although Reichardt and Mark (2001) concede that selection differences may be smaller when individuals are recruited from the same organization or locale. In conclusion, these authors advocate conducting *multiple analyses that consider a range of plausible assumptions* about the selection differences, but warn that even then, researchers need to be very cautious in interpreting the results from non-equivalent group designs. As Stevens (1999) summarised, "The fact is that inferring cause-effect from intact groups is treacherous, regardless of the type of statistical analysis. Therefore the task is to do the best we can and exercise considerable caution," (*p324*).

4.3.3 Addressing the Problem of Non-Equivalence

A number of approaches have been identified to address the problems described. These include adjusting the pretest scores for measurement error (Trochim, 2006) and proceeding with ANCOVA, although Reichardt and Mark (2001) point out that unless the pre-treatment measures have captured all of the selection differences that influence outcomes, the results of ANCOVA are likely to remain biased. They offer alternative solutions such as using matching and blocking techniques (although these require a large sample size) or change score analysis, (CSA). This involves calculating the difference between pre and post-scores for each individual; these 'change scores' thus become the dependent variable on which to perform a test to ascertain whether the 'mean change' between the groups is equal (Smolkowski, 2010). Fitzmaurice, Laird and Ware (2004) clarify how CSA and ANCOVA answer different research questions: the first addresses whether two *groups* differ in terms of their mean change; the second addresses whether an *individual* in one group is more likely to change than an individual in another group, "*given that they have the same baseline response*" (*p124*, emphasis in original).

There has been a historical debate about the reliability and validity of using change scores, as opposed to covariate analysis, which continues today (Rogosa and Willett, 1983; May and Hittner, 2010). Lord (1967, cited in Wright, 2006) illustrated how applying both approaches to the same data can lead to different results and conclusions; this so called 'paradox' can be a potential threat in intact group studies where individuals within groups change but the whole group does not (Maxwell and Delaney, 2004). As Wright (2006) notes, the approach adopted needs to take account of the research question, with change score analysis being preferred when the interest is in the *amount of gain for each group*.

Further criticisms include that the analysis of change scores can be questionable, particularly if the reliability of the measurement instrument is disputed (Gliner et al, 2003). Some have suggested that CSA can be particularly problematic with non-equivalent groups because if pre-test scores are unequal this may complicate the interpretation (see Smolkowski, 2010, for a review). Others have challenged this assumption, however; for example, Oakes and Feldman (2001) showed that even in the presence of baseline differences and measurement error, CSA yielded less bias and in some cases was more powerful than ANCOVA. This reflects Rogosa's (1988) demonstration of change score reliability and his conclusion that "The difference score is an unbiased estimate of true change" (*p180*). Further studies continue to provide evidence that CSA offers a more appropriate approach than partialling out the initial scores via ANCOVA for non-equivalent group designs (Fitzmaurice, 2001; Maxwell and Delaney, 2004).

The evidence presented here was considered sufficient to conduct a change score analysis in the present study.

4.3.4 Parametric versus Non-Parametric Analysis

The final consideration regards whether to apply parametric or non-parametric procedures in the following analysis. The issue of randomisation is central again to this decision as many researchers cite randomly allocated participants as a key assumption of parametric tests (Cohen et al, 2007; Gliner et al, 2003). However, others de-emphasise this requirement as long as data is normally distributed and displays homogeneity of variance (Dancey and Reidy, 2007; Coolican, 2009). Meanwhile, the other often cited requirement that data should be interval or ratio level (Coolican, 2009) has already been addressed. Dancey and Reidy (2007) add the final assumption that there should be no extreme scores as this can distort the mean upon which parametric tests are based.

Non-parametric tests, on the other hand, make no such assumptions about underlying population characteristics (Howitt and Cramer, 2011), and can thus be used in a wider variety of contexts (Robson, 2002), although Cohen et al (2007) suggest that certain non-parametric tests such as the Mann-Whitney U, still assume random sampling. The main cited disadvantage of using non-parametric measures is that they are typically less powerful than their parametric equivalents (Brace, Kemp and Snelgar, 2000), and parametric tests are potentially more helpful in detecting significant differences in smaller samples (Robson, 2002). Meanwhile, some have questioned the significance of violated assumptions in relation to parametric tests (Glass et al, 1972; Norman, 2010).

Unpicking these complexities, the researcher has opted to conduct both non-parametric and parametric tests, (where assumptions of normality are met), on the data in the present study. The rationale and method for this approach will now be presented.

4.4 Inferential Analysis in the Present Study

4.4.1 Method

The researcher has used Cohen et al's (2007) model as a framework for planning the current analysis (***Figure 4.9***).

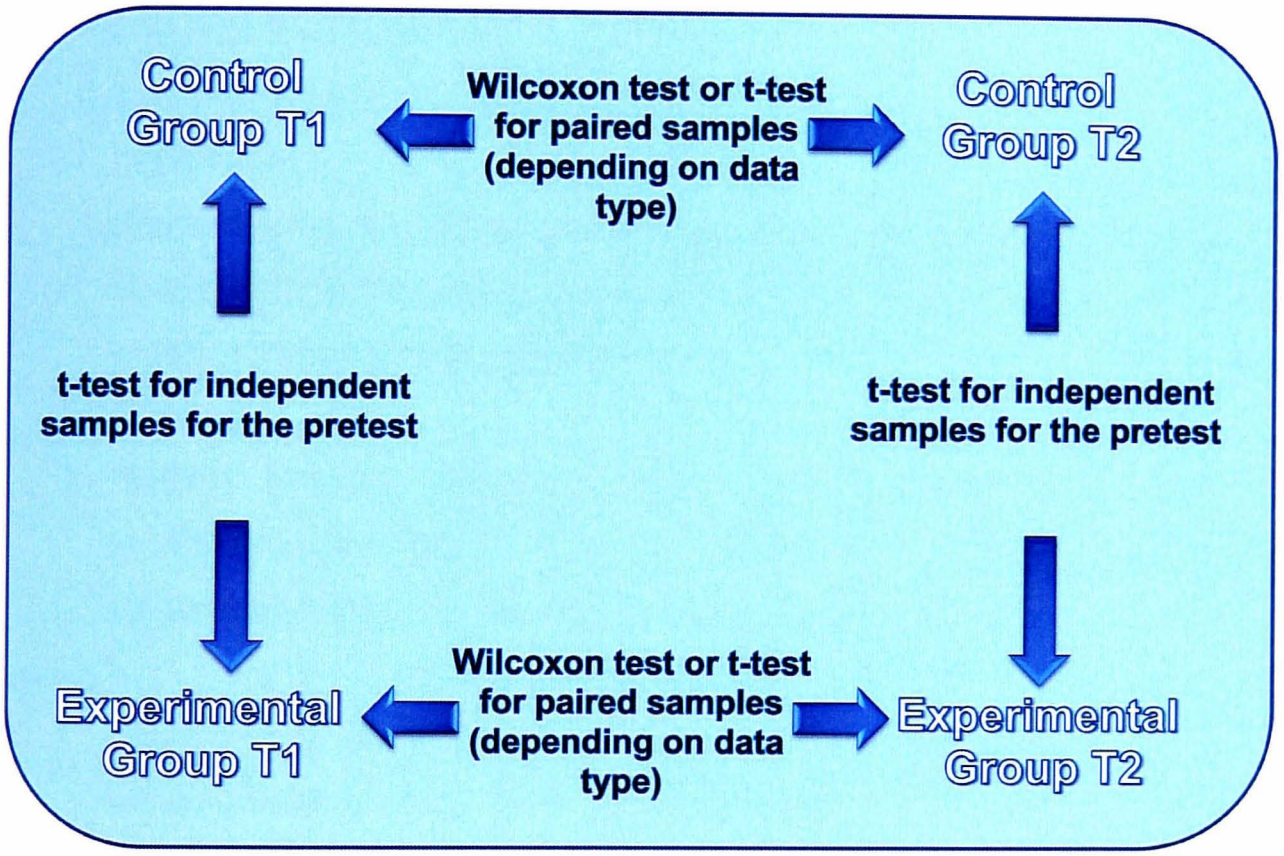


Figure 4.9: Identifying statistical tests for an experiment (Cohen et al, 2007, p587).

Acknowledging the limitations of the current design and method, the researcher has considered the following procedure to be the most comprehensive approach to addressing the research questions:

1. *Tests of skewness and kurtosis on the T1 and T2 data to indicate any areas of non-normal distribution in scores.* Normality was assessed according to the Shapiro-Wilk test; the null hypothesis of this test is that the sample is taken from a normal distribution, thus $p < 0.05$ for W rejects the supposition of normality (www.statsdirect.com). An

additional check involves calculating z scores by dividing the statistics for skewness and kurtosis by the standard error. If the z score lies below -1.96 or above 1.96 this indicates that the null hypothesis that the data is normally distributed cannot be rejected. (http://resources.esri.com).

2. *Independent Samples Mann-Whitney U tests of the T1 data* to indicate the equality of distribution of scores between the groups on each dependent variable at pre-test (Cohen et al, 2007). This highlighted whether any groups showed significant differences in scores before the intervention. A non-parametric test was selected due to some non-normal distribution revealed by the above analysis and, for the sake of consistency at this stage; it was considered appropriate to conduct the same procedure across all scores.

3. *Related Samples Wilcoxon Signed rank tests on every dependent variable for each group* to assess whether any significant change was observed between pre-test and post-test (Cohen et al, 2007).

4. *Independent Mann-Whitney U tests on the* ***change scores*** *for each group* (Gliner et al, 2007). This permitted an analysis of whether either group's distribution of change was significantly greater than the other's. Non-parametric procedures were selected initially as a conservative measure, to address the possibility of making a Type I error.

5. *Tests of normality and homogeneity of variance on the* ***change scores*** to assess whether any might meet the criteria for conducting parametric analyses (Dancey and Reidy, 2007).

6. *Independent samples t-tests on the* ***change scores*** that met the assumptions of normality and homogeneity of variance in order to address the possibility of making a Type II error (Robson, 2002).

The computer package PASW Statistics 18.0 has been used to conduct the inferential analyses (http://www.spss.com). The level of significance has been set at $p<0.05$ throughout this analysis as this is considered an acceptable threshold for supporting the research hypothesis or rejecting the null hypothesis (Dancey and Reidy, 2007, *p141*), although for reasons explained below, effect sizes will be reported alongside this statistic.

4.4.2 Effect Size

There is considerable debate in the social sciences regarding the application of statistical significance, (Dancey and Reidy, 2007; Cohen et al, 2007). Robson (2002) explains that if the *p* value is small rather than large, this makes it less likely that one's result is due to chance variation rather than a genuine difference, 'other things being equal.' This latter point is important, however, because as Coe (2000, cited in Cohen et al, 2007) points out the *p* value is unable to indicate whether it is the sample size or the coefficient, (ie. the size or importance of an effect or relationship), that is making the difference. Coe (2000) proposes that calculating the effect size addresses this deficiency and therefore can be interpreted as "measure of the effectiveness of the treatment" (*p1*).

There are various methods for calculating effect sizes (see Cohen et al, 2007, *p520-522*) but Pearson's *r* correlation has been selected for the present study as this can be applied to both parametric and non-parametric measures (Becker, 1999; Huber, 2011). Cohen (1988, cited on Cohen et al, 2007) proposed that effect size *r* may be interpreted as follows:

$r = 0.1 - 0.23$ (small effect)
$r = 0.24 - 0.36$ (medium effect)
$r > 0.37$ (large effect).

Effect sizes will be reported throughout this analysis, alongside significance levels, in order to help assess the strength of the impact of the intervention.

4.5 Results

4.5.1 Tests of normality

Selected outcomes from the tests of normality for both groups on the T1 and T2 data, illustrating examples of non-normal distribution can be found in Appendix 4d.

4.5.2 Tests of T1 distribution across both groups

Outcomes of this non-parametric analysis can be found in Appendix 4e. These indicated no significant discrepancies between the groups on any of the measures conducted at pre-test. This supports the view that scores were similarly distributed across both groups before the intervention.

4.5.3 Pre-test to post-test change for each group:

Wilcoxon Signed Rank tests yielded the following results (N = 38):

Dependent variable	Group	M (Median) pre and post		*Z*	*p*	Result	*r* (Effect size)
Emotional Distress (PI-ED)	IG	17.50	12.50	-2.16	.031	**Sig.**	.51
	CG	12.00	17.50	.393	.694	Not sig.	.09
Academic Self-Perceptions (MALS)	IG	68.00	68.50	1.45	.148	Not sig.	.34
	CG	65.00	65.00	.141	.888	Not sig.	.03
Total Difficulties (SDQ)	IG	10.00	6.00	-3.33	.001	**Sig**.	.78
	CG	12.00	8.50	-2.30	.021	**Sig.**	.51
Emotional Symptoms (SDQ)	IG	2.00	1.00	3.47	.001	**Sig.**	.82
	CG	3.00	1.00	-1.77	.076	Not sig.	.40
Conduct (SDQ)	IG	.50	1.00	-1.24	.214	Not sig.	.29
	CG	2.00	1.00	2.53	.011	**Sig.**	.57
Hyperactivity (SDQ)	IG	4.00	3.00	-2.59	.01	**Sig**.	.60
	CG	6.00	5.00	-.45	.654	Not sig.	.10
Peer Problems (SDQ)	IG	1.00	1.00	-1.95	.051	Not sig.	.46
	CG	2.00	1.00	-.43	.670	Not sig.	.10
Prosocial (SDQ)	IG	7.00	9.00	3.06	.002	**Sig.**	.72
	CG	5.50	7.50	2.58	.01	**Sig.**	.58

Table 4.16: Significance levels and effect sizes for pre to post test change for the IG and CG.

4.5.4 Change Score Analysis (Non-parametric)

Mann Whitney U tests yielded the following results (N=38):

Dependent variable	M (Median)		*U*	*p*	Result	*r* (Effect size)
	IG	CG				
Emotional Distress (PI-ED)	-3.00	1.00	246.0	.053	**Not Sig.**	.31
Academic Self-Perceptions (MALS)	4.00	-1.00	141.5	.260	Not sig.	.18
Total Difficulties (SDQ)	-4.50	-3.50	227.5	.163	Not sig.	.23
Emotional Symptoms (SDQ)	-1.00	-.50	216.0	.285	Not sig.	.17
Conduct (SDQ)	0.00	0.00	135.0	.119	Not sig.	.25
Hyperactivity (SDQ)	-1.50	0.00	247.0	.047	**Sig.**	.30
Peer Problems (SDQ)	5.00	4.00	227.5	.140	Not sig.	.24
Prosocial (SDQ)	1.00	1.00	149.5	.363	Not sig.	.15

Table 4.17: Significance levels and effect sizes for the non-parametric analysis of the change scores between both groups.

4.5.5 Change Score Analysis (Parametric)

Appendices 4f – 4m contain the results of tests of normality and homogeneity of variance across the change scores. This analysis indicated that the assumptions for parametric tests outlined in section were fully met by the following sets of scores:

- Emotional Distress
- Academic Self-perceptions
- Total Difficulties.

All of the SDQ subscale scores, except Hyperactivity, failed to meet the assumptions of normal distribution and were therefore excluded from the parametric analysis. A parametric analysis of the Hyperactivity scores was conducted with the 'equal variances not assumed' value reported.

Pertinent PASW output may be found in Appendices 4n – 4q.

Independent samples *t*-tests yielded the following results:

Dependent variable	**M (Mean) /SD**		***T* (*df* = 36)**	***p***	**Result**	***r* (Effect size)**
	IG	**CG**				
Emotional Distress (PI-ED)	-4.67 (8.09)	.75 (6.89)	-2.23	.032	**Sig.**	.34
Academic Self-Perceptions (MALS)	5.83 (13.91)	-.15 (10.29)	1.52	.138	Not sig.	.24
Total Difficulties (SDQ)	-4.67 (4.31)	-2.40 (3.97)	-1.69	.100	Not sig.	.26
Hyperactivity (SDQ)	-1.89 (2.63)	-.15 (1.53)	-2.45	.021	**Sig.**	.37

Table 4.18: Significance levels and effect sizes for the parametric analysis of the change scores between both groups.

4.6 Summary of Findings

This chapter has highlighted some limitations with regard to the design, measures and analysis in the present study. These include the non-randomisation of participants, limited sample size and the use of ordinal data that has restricted the kinds of statistical analysis that can be applied (Gliner et al, 2003). However, the researcher has attempted to carry out as thorough an analysis as possible, including exploring the normality of each data set, calculating effect sizes alongside significance values and conducting the more robust parametric analyses where assumptions have been appropriately met.

The following conclusions have been drawn in relation to the research questions:

4.6.1 Does a class of Key Stage 2 children participating in a universal FRIENDS intervention report a significant reduction in emotional distress (ED) in comparison to a non-intervention control group?

Experimental Hypothesis: A class participating in a universal FRIENDS intervention will report a significant reduction in ED compared to a non-intervention control group.

Null hypothesis: There will be no significant difference in ED reported by a class attending a universal FRIENDS intervention and a non-intervention control group.

Summary: Wilcoxon Signed Ranks tests indicated that ED as reported on the PI-ED reduced significantly for the IG between pre-test and post-test (Z = -2.157, p = .031, r = .51), but not for the CG (Z = .393, p = .694, r = .09).

A Mann-Whitney test on the ED change scores indicated that there was no significant difference in changes in ED between the IG (M = -3.00) and the CG (M = 1.00), U = 246.00, p = 0.053, r = 0.31. However, an independent

samples t-test indicated that emotional distress scores for the IG (M = -4.67, SD = 8.09) decreased significantly more than those for the CG (M = 0.75, SD = 6.89), $t(36)$ = - 2.23, p = .032, r = .34.

Wilcoxon Signed Ranks tests indicated that teacher-reported emotional symptoms on the SDQ decreased significantly between pre-test and post-test for the IG (M = 2.Z = 3.47, p = .001, r = .82) but not the CG (Z = -1.772, p = .076, r = .40).

A Mann-Whitney test on the emotional symptoms change scores indicated that there was no significant difference in changes in emotional symptoms between the IG (M = -1.00) and the CG (M = -0.50), U = 216, p = .285, r = 0.17.

Conclusion: The results from the PI-ED indicate some support for the experimental hypothesis. Although the non-parametric analysis of change scores did not achieve statistical significance this was only missed by a small degree and the other analyses revealed trends in reduced ED favouring the IG. The results from the teacher-rated emotional symptoms change scores indicated only a weak effect, although when considered separately, the IG's emotional symptoms scores decreased significantly according to teacher report, whereas the CG's did not.

Taken together, the experimental hypothesis is supported by the self-reported emotional distress scores and *partially* supported by the teacher-reported scores.

4.6.2 Does a class of Key Stage 2 children participating in a universal FRIENDS intervention experience a preventive effect for ED in comparison to a non-intervention control?

Experimental Hypothesis: Reports from a class participating in a universal FRIENDS intervention will indicate a preventive effect for ED in comparison to those in a non-intervention control group.

Null hypothesis: There will be no difference in reported ED preventive effects between a class attending a universal FRIENDS intervention and a non-intervention control group.

Summary: The number of children in the IG experiencing elevated levels of emotional distress reduced from around half of the class to a third at post-test. Meanwhile, the CG showed the opposite trend with a third of the class showing elevated levels at pre-test, moving to over a half at post-test. Of the ten children whose scores were originally 'at risk' in the IG, five moved out of that range and of the five that stayed in, four demonstrated reduced scores at post-test. In contrast, for all of the children whose scores started in the 'at risk' range in the CG remained there at post-test, and of these seven, five had increased scores at post-test. Four further children's scores in the CG moved into the 'at risk' range, whereas only one in the IG tipped into this range with a score of 16.

Conclusion: As more children in the IG moved out of the range of elevated distress and fewer children moved in, whereas the CG showed trends in the opposite direction, this provides support for the experimental hypothesis that the IG children's reports would indicate a preventive effect for the intervention in comparison to the CG.

4.6.3 Does a class of Key Stage 2 children participating in a universal FRIENDS intervention display significantly more positive academic self-perceptions than those in a non-intervention control group?

Experimental Hypothesis: A class participating in a universal FRIENDS intervention will report a significant improvement in academic self-perceptions in comparison to a non-intervention control group.

Null hypothesis: There will be no significant difference in academic self-perceptions reported by a class attending a universal FRIENDS intervention and a non-intervention control group.

Summary: Wilcoxon Signed Ranks tests indicated that there was no significant difference in academic self-perceptions between pre-test and post-test for the IG (Z = 1.45, p = .148, r = .34) or the CG (Z = .141, p = .888, r = .03).

A Mann-Whitney test indicated that there was no significant difference in changes in academic self-perceptions between the IG (M = -3.00) and the CG (M = 1.00), U = 141.50, p = .260, r = .18 and this was confirmed by an independent t-test, IG (M = 5.83 , SD = 13.91); CG (M = -.15, SD = 10.29), $t(36)$ = 1.52, p = .138, r = .24.

Scrutiny of the children moving in and out of the 'low academic self-perception range' indicated that around a third of the IG started in this range at pre-test, but this reduced to just 5.5% at post-test. The CG started with around a quarter of the class in the low range at pre-test, rising to over a third at post-test. Of the original six children in the IG with low academic self –perception scores at pre-test, five had raised their scores to within the average range at post-test and, although one child remained with a deflated score, no further children moved into the below average range. In the CG, however, only one child's score rose to within the average range, four remained below average and a further three fell to below average.

Conclusion: Taken together, the inferential analyses upheld the null hypothesis that no significant differences would be observed in self-reported academic self-perceptions between the groups. However, the 'preventive effects' analysis indicates more positive trends in relation to low academic self-concept for the IG than the CG.

4.6.4 Does participation in a universal FRIENDS programme result in a significant improvement in teacher-rated pupil behaviour (reduced difficulties and increased prosocial behaviour) in comparison to a non-intervention control?

Experimental Hypothesis: Teachers of children participating in a universal FRIENDS intervention will report significantly improved behaviour (reduced difficulties and increased prosocial scores) in comparison to those in a non-intervention control group.

Null hypothesis: There will be no significant difference in teacher reports of behaviour (reduced difficulties and increased prosocial scores) for those attending a universal FRIENDS intervention and those in a non-intervention control group.

Summary: A Wilcoxon Signed Ranks test indicated that total difficulties decreased significantly between pre-test and post-test for both the IG (Z = -3.33, p = .001, r = .78) and the CG (Z = -2.30, p = .021, r = .51).

A Mann-Whitney test indicated that there was no significant difference in changes in total difficulties between the IG (M = -4.50) and the CG (M = -3.50), U = 227.5, p = .163, r = 0.23 and this was confirmed by an independent t-test: IG (M = 5.83 , SD = 13.91); CG (M = -.15, SD = 10.29), $t(36)$ = 1.52, p = .138, r = .24.

A Wilcoxon Signed Ranks test indicated that there was a significant increase in prosocial behaviour between pre and post-test for both the IG ($Z = 3.062$, $p = .002$, $r = .72$) and the CG ($Z = 2.58$, $p = .01$, $r = .58$).

A Mann-Whitney test indicated that the difference in prosocial behaviour changes was not significant between the IG ($M = 1.00$) and the CG ($M = 1.00$), $U = 149.50$, $p = .363$, $r = .15$.

Results for the other disaggregated SDQ scores (Conduct, Hyperactivity, Peer Problems) will be discussed further in Chapter Five when considering the possible processes working in this intervention.

Conclusion: The SDQ results suggest some interesting (and, in places, significant) trends, which will be discussed in more detail in the next chapter, but the null hypothesis is upheld that no overall significant differences between the groups in behaviour change would be observed.

Chapter Five explores these findings, the relationships between them and their potential implications in further depth.

4.7 Post Hoc Correlational Analyses

In order to clarify some of the questions arising in the Discussion, the researcher conducted a number of 'post hoc' correlations on the following change scores to see whether certain constructs were linked for this sample of children: emotional distress and emotional symptoms; emotional distress and academic self-perceptions; hyperactivity and emotional distress; hyperactivity and prosocial; emotional distress and prosocial.

The results from this analysis can be found in Appendices 4r – 4v and are discussed in Chapter Five.

5 CHAPTER FIVE: DISCUSSION

Introduction

In this section, the outcomes and conclusions from the previous chapter are considered in more depth in order to evaluate both their internal and external validity. Each research question is addressed independently with reference to the theoretical links from Chapter Two, to explore how these findings correspond with expectations from previous literature. The researcher also makes reference to the methodological issues that were highlighted in Chapter Three.

The discussion then broadens to a consideration of how this study extends understanding in relation to previous FRIENDS research. The observation from cross-cultural applications that adapting the programme to fit context may be implicated in effecting change (see section 2.5.4) is given particular attention. This involves revisiting the critical realist and social cognitive perspectives to consider the role of contextual factors and possible 'mechanisms' that may have contributed to the study outcomes (Kazdin and Nock, 2003). Drawing on observations from the present study, the researcher considers how the findings contribute to the wider literature about the current role of EPs supporting mental health initiatives in schools.

5.1 Research Question 1:

Does a class of Key Stage 2 children participating in a universal FRIENDS intervention report a significant reduction in emotional distress in comparison to a non-intervention control group?

Emotional distress was defined as comprising anxiety and depression and was assessed using the Paediatric Index of Emotional Distress (O'Connor et al, 2010) and the teacher-rated Emotional Symptoms subscale on the SDQ. It was hypothesised that the components of the FRIENDS programme,

addressing cognitive, behavioural and physiological factors, would facilitate significantly reduced levels of emotional distress for the IG compared to the CG.

5.1.1 PI-ED

Scrutiny of the descriptive data for the PI-ED indicated that both mean and median emotional distress scores for the IG decreased while those for the CG slightly increased. Related samples non-parametric tests indicated that the IG's post-test scores were significantly lower than their pre-test scores (p = .031) with a large effect size (r = .51), while the CG's showed no significant difference. When the change scores were analysed, the non-parametric test just missed the selected significance value (p= .053); however, the effect size fell within the moderate range, (r = .31). As this data met the assumptions for parametric testing a t-test was conducted which indicated that the IG demonstrated a significant reduction in emotional distress compared to the CG (p = .032), with an effect size in the high moderate range (r = .34). Taken together, these results provide some support for the hypothesis that the IG would report a significant reduction in emotional distress in comparison to the non-intervention control.

Threats to Validity

A number of confounding variables represent threats to the validity of this conclusion, however. As was expounded in section 3.16, an important issue in educational evaluations using this design is the inability of the researcher to *blind participants* to the nature of the study (Mertens, 2010). In this context, the children were conscious that they were participating in a programme to raise their awareness about emotions and behaviours. It is therefore likely that both the Hawthorne effect and 'testing' alerted them to the constructs under scrutiny (Cohen et al, 2007). This may have affected their post-test responses, possibly encouraging them to respond in a way that they perceived as being socially desirable or sought after by the researcher (Lewis and Lindsay, 2000). It is difficult to refute these possibilities without

triangulating the evidence with some of the additional assessment methods suggested by Lewis and Lindsay (2000), for example. However, when the researcher informally questioned the children following the T2 test, to see whether any thought they had answered differently that time, there was not an apparent awareness of the anticipated direction of change from the children's responses.

Another claim could be that the differences in change scores may be attributed to *selection-by-maturation interaction* (Robson, 2002). This describes how treatment groups change at different rates so that gaps between them reduce or widen over time, even in the absence of treatment effects (Reichardt and Mark, 2001). A further threat to the present analysis involves the IG's particularly high score at pre-test, decreasing to a lower score, which may be interpreted as a *regression* effect (Robson, 2002). There are two important responses to these challenges. The first involves the non-parametric analysis, which revealed that, despite appearing inflated, the distribution of the IG's pre-test scores was not significantly different to the CG's, indicating that the groups were similar at pre-test. Secondly, Figure 4.1 in Chapter Four displays a crossover interaction between the IG's and CG's mean scores. (This describes when the IG's mean score starts above the CG but ends up lower; Reichardt and Mark, 2001). Reichardt and Mark (2001) suggest, "such a pattern seldom can be plausibly explained as due to either a selection-by-maturation interaction or regression towards the mean" (*p10656*). This conclusion is supported by the outcomes of the prevention effects analysis, which appear to be considerably stronger for the IG, and it is difficult to explain how these differences would have emerged through the natural course of development. This suggests that at some level, the IG's levels of emotional distress were affected by the intervention, although the precise mechanisms underlying this influence require further consideration (see Section 5.5).

Cohen et al (2007) point to the validity of the measurement instrument itself when analysing the ability to generalise conclusions. It was reported in Chapter Three that the PI-ED is a new measure, on which the initial

standardisation is being reviewed and enhanced (Personal communication with E. Ferguson, 01.04.11). The issue of the clinical cut-off points has already been highlighted. Further concerns were raised during a consultation between the researcher and one of the test authors, regarding the stability of the measure and the wording of its items. Although the test requires children to "Think about how [they] have been feeling over the last week," the items are phrased in the present tense, e.g. "I feel happy". This has alerted the author to consider whether the measure is in fact, assessing a relatively stable trait such as anxiety or a more dynamic state such as 'mood.' Meanwhile, the proportion of children in this study from non-White UK backgrounds was much higher than the sample on which the PI-ED was standardised, and thus cultural factors may be implicated in the elevated scores (Barrett, 2000). While these issues are under review, the researcher would highlight the compatibility between the present results and those of other FRIENDS studies employing more established anxiety and depression measures such as the SCAS (Spence, 1994) and the CDI, (Kovacs, 1981). Furthermore, the study by Lau et al (2008) cited in Chapter Two, revealed the strong influence of environmental factors on *state* anxiety. Thus, even if the measure's assessment of trait anxiety is questionable, these results still uphold that the IG's state of emotional distress was lower, possibly due to situational adaptations since the beginning of the intervention. This point will be revisited in section 5.6.

5.1.2 SDQ: Emotional Symptoms

This subscale included consideration of whether the children exhibited worries, nervous or fearful behaviours, for example.

Descriptive analysis indicated that mean emotional symptoms scores decreased for both groups between pre and post-test. Non-parametric analysis showed that emotional symptoms reduced significantly for the IG (p = .001), but not for the CG (p = .076). The effect for the IG (r = .82) was much larger than for the CG (r = .40) although both were in the high range. Non-parametric analysis of the change scores indicated that there was no

significant difference in changes in emotional symptoms between the IG and the CG (p = .285). Taken together, these results suggest that teacher-rated emotional distress for the IG decreased significantly over the period of study, but not significantly more than the CG's and thus the reduction cannot be attributed reliably to the intervention.

It may be argued that sensitization effects influenced the decrease in emotional symptoms scores because the teachers were aware of the nature and aims of the intervention (Cohen et al, 2007). Research has also highlighted poor agreement between self-report and proxy-report in emotional domains on certain measures (Limbers et al, 2008). Correlational analyses indeed revealed that the association between the pupil and teacher reports was very weak (r = -.021, p = .901), indicating that even though both reported significant differences, these were not necessarily applicable to the same individuals. This might also echo suggestions that the SDQ may be insensitive to detecting important changes in anxiety symptoms (Stallard, 2010).

5.1.3 Summary

The perils of implying causal inference from quasi-experimental designs have been emphasised in Chapter Four and highlighted by discrepancies in this analysis. Broadly, the outcomes of this study support those of the other British and Australian studies with regard to the impact of FRIENDS on levels of emotional distress, (refer to section 2.5.4, summary), although the self-report measures are more pronounced than the teacher-rated scores. The wider implication for this analysis for this domain of research and for EP practice will be discussed later.

This section has focused on the group reduction of emotional distress scores as rated by pupils and teachers. The next question addressed whether children in the IG were more likely to move out of and/or less likely to move into the higher range scores on the emotional distress continuum.

5.2 Research Question 2:

Does a class of Key Stage 2 children participating in a universal FRIENDS intervention experience a preventive effect for emotional distress in comparison to a non-intervention control?

For this question, emotional distress was measured by self-report on the PI-ED. It was hypothesised that children in the IG would experience a greater preventive effect for emotional distress than those in the CG.

The difficulties associated with applying clinical cut-off points were described in section 4.1.2, while authors of the PI-ED are analysing further data from a wider variety of samples to re-calculate sensitivity and specificity thresholds. Following advice from one author, the researcher adjusted the level of clinical cut-off for the purposes of analysing prevention effects to 15. (It was noted that if the threshold was raised further by another couple of points, then the pattern of results remained relatively stable). Considering the higher estimates of anxiety prevalence rates reported in section 1.1, (10-20%, Barrett and Pahl, 2006), it was observed that a particularly large number of children fell within the elevated range for this sample (IG: 55.5%; CG: 35%). It is therefore important to recall Connell et al's (2007) caution regarding the context and purpose of cut-off scores. Indeed, a social cognitive perspective would refute the idea of clinical cut-offs, advocating psychological continuity between problem and non-problem responses (Williams, 2005). Thus, the figure is considered more arbitrary in the present analysis rather than indicating the presence of 'symptoms' requiring individual intervention.

As was seen in sections 4.1.2. and 4.6.2, a greater percentage of children in the IG moved out of the elevated range of emotional distress and a smaller percentage moved in (this represented one child who scored close to the cut-off). The total percentage of the class in this range decreased in the IG from over half at pre-test to around a third at post-test, whereas the CG showed the opposite trend. Indeed, nearly all of the 'at risk' children in the IG showed a reduction in scores at post-test, which was not matched by the CG.

5.2.1 Summary

Although it was not deemed appropriate to undertake a statistical analysis of this data, (Mertens, 2010), taken with the overall results for emotional distress, these observations provide tentative evidence for a preventive effect corresponding with the intervention. These results mirror those of Stallard et al (2008) which also indicated a reduction in the number of children classified as 'high risk' at post-test and no 'low risk' children' moving into this group, which lends further support for the preventive application of FRIENDS in universal settings.

5.3 Research Question 3:

Does a class of Key Stage 2 children participating in a universal FRIENDS intervention display significantly more positive academic self-perceptions than those in a non-intervention control group?

Academic self-perceptions were defined as comprising the key constructs of academic self-concept and academic self-efficacy and were measured using the Myself-As-Learner Scale (Burden, 1998). It was hypothesised that by reducing levels of emotional distress and promoting coping strategies, there would be an associated increase in the IG's academic self-perceptions that would be significantly greater than a non-intervention control. The rationale for this link was presented in Chapter Two.

The descriptive analysis indicated that the IG's mean score showed a slight rise in academic self-perceptions at post-test, while the CG's remained relatively stable. It was noted that one extreme score was observed in the IG's post-test scores, which may have skewed the mean. Inferential analyses revealed that there was no significant difference in academic self-perceptions for the IG (p = .148) or the CG (p = .888) between pre and post-test. Parametric analysis of the change scores for both groups confirmed that there

was no significant difference in changes in academic self-perceptions between the IG and the CG (p = .10) with a low medium effect (r = .24). However, independently, the effect size was revealed to be stronger for the IG (r = .34) than the CG (r = .03). Even when the extreme score from participant 12 was excluded from the analysis (Dancey and Reidy, 2007), the effect size for the IG remained in the moderate range (r = .30). Overall, despite the positive trends in effect size, these results support the null hypothesis that little significant change in academic self-perceptions would occur between the groups as a result of the intervention. There are a number of potential explanations for why a stronger effect for the intervention on academic self-perceptions was not demonstrated in this study and these will be explored below.

5.3.1 Timescale

One hypothesis is that the timescale did not allow for significant improvements in academic self-perceptions to be demonstrated. This would be supported by Usher and Pajares's (2008) finding that mastery experience is the most influential source of self-efficacy beliefs, implying that the children in the IG would need time to practise and apply their newly acquired coping skills before demonstrating improved self-efficacy to a more significant level. A follow-up measure several months after the intervention would enable this hypothesis to be explored.

5.3.2 Absence of links between emotional distress and academic self-efficacy

Secondly, it is possible that the espoused links between reduced emotional distress and improved academic self-efficacy described in section 2.3.7 were not evident in this sample. However, a Pearson's correlational analysis of the change scores for the PI-ED and MALS revealed *a significant association between reduced emotional distress and increased academic self-perceptions* across the year group (r = -.693, p = 0.00). This lends validity to the measures employed because it reflects the assumptions of self-efficacy

theory and the outcomes of previous studies, (refer to section 2.3.7). Self-efficacy theory predicts that strengthened coping efficacy leads to lowered anxiety and less avoidant behaviour (Bandura 1988). The correlational analysis of change scores across the year group supports such an association between increased academic self-efficacy, including improved confidence in learning and problem-solving, and lowered emotional distress. However, it was not feasible to ascertain the direction of causal influences from this analysis and indeed, additional variables may have influenced both constructs separately. (As no associations were detected between individual SDQ variables and both academic-self-perceptions and emotional distress, however, it would appear that this theory remains outside the scope of the current analysis). Furthermore, although the correlation was stronger for the IG ($r(18) = -.76$, $p = .000$) than the CG ($r(20) = -.55$, $p = .012$), as the latter group also showed a moderate correlation between these two variables it cannot be assumed that the intervention was implicated in any *associated change* between the two. To summarise, although some anticipated links were found, the relationship between the intervention, improved academic self-efficacy and reduced emotional distress as predicted by SCT was not seemingly observed. However, it is possible that this result reflects methodological rather than theoretical limitations and this will now be discussed further.

5.3.3 The dependent variable was not adequately operationalised

Cohen et al (2007) raised the issue of whether the methods employed accurately reflect the constructs under scrutiny. Although evidence for the construct validity of the MALS has been provided in section 3.10.2 and in the section above, it is possible that it did not effectively capture individual improvements in self-concept and self-efficacy related to particular domains of academic functioning, and was only successful in capturing some broad trends. This reflects the difficulties of operationalising abstract constructs related to the self described in section 2.4.2. Burden (1998a) described the measure's intention to "de-emphasize the structural nature" of self-perceptions and to present them as "transitory, though connected and open to

change by means of appropriate intervention," (*p296*). However, given the support for the structural nature of academic self-concept, (Marsh, 1990; Bong and Skaalvik, 2003), it must be questioned whether the measure was precise enough to detect subtle changes in academic self-perceptions. The robust observation that assessment must be pitched at the appropriate level of specificity is salient here (Usher and Pajares, 2008). According to Marsh and Shavelson's (1976) model, 'academic self concept' is in a central and relatively stable position in the hierarchy and thus would not be as easily influenced by a short-term intervention as subject-specific self-concept (Marsh, 1990). Indeed, anecdotal evidence from school staff indicated that individuals *were* applying the FRIENDS strategies (such as 'coping step plans' and 'green thoughts') to problematic areas of study and this was having a positive effect on their engagement, although this has not been effectively evidenced by the present methodology. Thus, for children using their coping step plans to develop confidence in spelling, for example, evidence suggests that a scale to measure 'self-perceptions in spelling' is more likely to have detected change. The opportunity to assess the impact of the intervention on academic self-perceptions in individual subject areas has therefore not been addressed and may be a focus for future study.

5.3.4 Support for FRIENDS promoting a healthy academic self-concept

Scrutiny of the children moving in and out of the 'low academic self concept' range provides more apparent support for the effects of the intervention (see sections 4.1.2 and 4.6.3).

While the validity of using cut-off points has been questioned and the MALS data was standardised on an older sample of pupils, these results do indicate some interesting trends in line with the research hypothesis. Furthermore, scrutiny of the individual question data showed that the IG increased their scores on all of the items loaded onto the 'self-efficacy' factors of the scale, except Item 1 ("I'm good at doing tests"), in which both groups went down. This is perhaps reminiscent of Carr's (2006) point that test anxiety becomes particularly salient during the middle childhood years and reinforces the

importance of addressing this. An interesting observation was that the IG made two particularly large 'leaps' between T1 and T2 on item 10, "When I am stuck with my work I can usually work out what to do next" and item 15, "I know how to solve the problems that I meet," whereas the CG's total scores decreased on these items. Both of these items reflected key teaching points from the programme and so the results could reflect genuine learning or possibly, the children responding in the socially desirable direction (Lewis and Lindsay, 2000).

5.3.5 Summary

There is insufficient evidence to suggest that the intervention had a direct result on improving academic self-perceptions. However, the strength of the effect of the intervention on the IG was shown to be stronger than for the CG and analysis of preventive effects for low academic self-concept also favoured the IG. This evidence might indicate that children taught proactive coping skills, including problem-solving and cognitive reframing, in relation to academic difficulties, might be less likely to use avoidant and non-productive strategies and therefore feel more able to tackle the problems that they meet (Zimmer-Gembeck and Skinner, 2011; Frydenberg, 2008). Furthermore, a number of methodological limitations have been described which may have affected the accurate measurement of academic self-perceptions (Usher and Pajares, 2008). On reflection, perhaps the research question could have been adjusted to "academic self-perceptions in targeted areas of learning," although this would probably have necessitated a more individualized methodological approach such as Single Case Experimental Design (Barlow, Nock and Hersen, 2008).

5.4 Research Question 4:

Does participation in a universal FRIENDS programme result in a significant improvement in teacher-rated pupil behaviour (reduced difficulties and increased prosocial behaviour) in comparison to a non-intervention control?

In this section, the results from the total difficulties and individual subscales on the SDQ are discussed to explore any significant relationships involving the discrete constructs (conduct, hyperactivity, peer problems, prosocial behaviour) which might inform understanding of additional processes working in this context. It was hypothesised that the children in the IG would show a significantly greater improvement in overall teacher rated behaviour than those in the CG.

5.4.1 Total Difficulties

The subscale scores form the SDQ (excluding Prosocial) were aggregated to give a Total Difficulties score. Non-parametric analyses revealed that this reduction in teacher-rated Total Difficulties was significant for both the IG (p = .001, r = .78) and the CG (p = .021, r = .51) and parametric change score analysis confirmed that the differences between the groups was not significant, (p = .100, r = .26).

These results indicate that the significant overall effects on the children's behaviour reported in the IG cannot be attributed solely to the intervention. The possible explanations for this occurrence include changes in the participants' environment unrelated to the programme or maturation effects across the year group (Robson, 2002). Equally, the intervention *may* have influenced the IG's behaviour, and improvements were 'diffused' to the CG (Robson, 2002). Alternatively, the T2 measures may have given a more 'accurate' profile of the children's behaviour than the T1 scores because the teachers' knowledge of the children was more thorough by the spring term. Further analysis of the subscales reveals some interesting differences

between the groups in the specific areas of improvement, however, some of which may be attributable to the intervention, and these will now be discussed.

5.4.2 Conduct

This subscale comprised observations about the child's externalizing behaviour, including temper tantrums and general levels of obedience. Recalling the literature on comorbidity in section 2.2.5, the Conduct and Hyperactivity subscales have been included to explore any corresponding changes in externalizing behaviours alongside emotional distress.

Descriptive analysis revealed that behaviour problems reduced for both the IG and the CG between pre and post-test. Inferential analyses revealed that the CG's behaviour problems reduced significantly (p = .011) but the IG's did not (p = .214). Change score analysis indicated no significant difference in the rate of behaviour change between the groups (p = .119)

While these results are in a socially desirable direction, the evidence does not support a relationship between the intervention and reduced conduct problems. It is likely that the changes in the CG were therefore due to history or maturation variables, or possibly regression effects as this group started with a higher score, (Robson, 2002). These results reflect the inconsistent results for the effects of CBT on externalizing behaviours (Wolpert et al, 2006).

5.4.3 Hyperactivity

This subscale focused on issues related to restlessness, concentration span and impulsivity, for example.

Descriptive analyses indicated that both groups showed a decrease in teacher-rated hyperactivity scores. Non-parametric analyses of the change scores indicated that the IG showed a greater reduction in hyperactivity than

the CG (p = .047, r = .30) and this was strengthened in the parametric analysis (p = .021, r = .37).

Given the non-linear relationship between anxiety and hyperactivity/inattention described in section 2.2.5 and the inconclusive evidence for the effect of CBT on attention problems, these results are intriguing. One hypothesis remains that the programme components facilitated a reduction in emotional distress with a corresponding impact on hyperactivity – or vice versa. (A Pearson's correlation confirmed that the association between the hyperactivity and emotional distress change scores was weak, however, r = .20). A further suggestion is that the intervention has affected a *third variable* such as "negative affectivity," (defined in section 2.2.5), which has impacted differentially on levels of both emotional distress and hyperactivity. A third possibility relates to the fact that the relaxation component of the programme was given extra emphasis and reinforced through additional Relax Kids extracts (www.relaxkids.com). However, as the CG also had access to Relax Kids, it is unlikely that this is the sole explanation for the reduction in hyperactivity scores, and therefore the role of additional programme components (such as cognitive restructuring, coping and problem-solving skills), should also be considered. Another explanation is that the *psychological experience of participating* in an intervention promoting self-regulatory behaviour, (rather than the intervention itself), produced lower levels of observed emotional distress and hyperactivity, and/or possibly affected the teacher's perception of these constructs ('Hawthorne effect,' Cohen et al, 2007).

Adopting a realist perspective, the contribution of social and environmental influences must also be acknowledged. For example, evidence suggests that children judged to have hyperactivity may be more likely to have conflicts with adults and peers, and suffer from unpopularity and a lack of friendships (Nijmeijer et al, 2008). It is therefore possible that the improvement in social relations witnessed across the year group (see sub-section *Prosocial Behaviour* below) was a key factor in influencing hyperactivity levels *and* emotional distress. A Pearson's correlational analysis of the hyperactivity and

prosocial change scores indicated that this association was indeed significant, (r = -.36, p = .026), although emotional distress and prosocial change were unrelated, (p = .902). Applying SCT, it is possible that a self-report measure of self-efficacy in friendships/social relations would illuminate the nature of these associations further.

In summary, these results suggest that participating in FRIENDS may *contribute* to lowered emotional distress and hyperactivity as measured by the PI-ED and SDQ, although it is questionable whether these phenomena are directly related. The influence of possible confounding or mediating variables such as the Hawthorne effect and prosocial influences has been highlighted, while the need to triangulate these results with additional reports or methods (eg. observation) is apparent (Lewis and Lindsay, 2000). It is therefore not possible, given the limitations of the current design, to conclude a *causal relationship* between FRIENDS and reduced hyperactivity.

5.4.4 Peer Problems

This subscale focused on the quality of the relationship between individuals and their peers and adults in school.

Descriptive analyses revealed that both the IG and the CG showed reductions in peer problems and change score analysis indicated that differences between the groups were not significant (p = .140, r = .24). However, independently, the IG's reduction just missed statistical significance (p = .051) with an effect size in the large range, (r = .46) in comparison with the CG (r = .10). According to anecdotal evidence from the class teacher, the children in the IG *were* demonstrating considerably increased co-operation with their peers both in the classroom and at playtime since the commencement of the intervention. This was a hypothesised outcome due to the programme's emphasis on peer learning and the joint sharing of experiences (Barrett, 2004). It is possible, therefore, that because scores were within the low range to begin with, this measure has not been comprehensive enough to detect situation-specific changes in classroom behaviour. (This represents a

'floor effect', where performance scores are very low or absent at pre-test, Kantowitz, Roediger and Elmes, 2009, *p294*).

5.4.5 Prosocial Behaviour

This subscale focused on skills such as empathy and altruistic behaviour.

Inferential analyses revealed a significant change in teacher-rated prosocial behaviour for both the IG (p = .002) and the CG (p = .01) with large effect sizes for both groups, (r = .72 and r = .58 respectively). Change score analysis indicated that there was no significant difference in the rate of change between the groups (p = .363).

Although this supports the null hypothesis that no significant difference would be observed *between* the two groups, the significant changes *within* them are worthy of comment. It is possible, for example, that the IG's improvement is intervention-related, as would be supported by the 'Peer Problems' scores, and a *diffusion effect* has occurred across the groups (Robson, 2002). As this variable is the most likely to be played out during joint times, (such as playtime) it is possible that prosocial changes in the IG have spilled over into their relationships with the CG. This supports increasing evidence for the effects of the programme on widening friendship groupings (Stallard et al, in press). Alternatively, it may be that history effects across the groups were implicated in this change and the intervention had little or no effect upon this variable (Robson, 2002). In the absence of further evidence to qualify the sources of change, it is impossible to be certain, although this result invites exploration in future studies.

5.4.6 Summary

The IG showed significant reductions in Emotional Symptoms and the CG showed significant reductions in Conduct problems. Teacher-reported reductions in Hyperactivity were significantly greater for the IG than the CG. Both groups showed significant improvements in Total Difficulties and

Prosocial behaviour. Overall, the IG's reduction in Total Difficulties was not significantly greater than the CG's, which suggests that *IG improvements cannot be reliably attributed to FRIENDS*. Alternatively, the CG's teachers may have inflated their post-test scores, reflecting 'compensatory' validity threats (Robson, 2002). Not only do these uncertainties reflect the importance of including a control group in the design to highlight potential confounding effects (Stallard, 2005), but they also reinforce the complexity of attempting to unravel in real world research "what it is that is changing in development" (Dockrell, Lewis and Lindsay, 2000). One conclusion might be that the intervention formed part of a web of influences that contributed to the significant changes reported above (Morrison, 2002). If this is the case, it would be important to explore some of the mechanisms that might have been operating (Kazdin, 2008), in order to address the purposes outlined in Chapter One: to identify the context specific factors that facilitated implementation of the programme for the project school and to consider which ones might be helpfully generalised to other universal applications.

5.5 Evaluating the Impact of FRIENDS: A Realist Interpretation

As has been emphasised throughout this analysis, establishing causal inference from a quasi-experimental design in a real-world setting is hazardous (Stevens, 1999), and the aim of this section is not so much to uncover scientific 'truths,' but to generate discussion from theoretical insights about how the intervention might be 'fine-tuned' to maximise its impact.

The term 'mechanisms' refers to the processes or events that lead to therapeutic change (Kazdin and Nock, 2003). In psychotherapy research, a distinction has been drawn between 'mediators' (the cause or mechanism of change) and 'moderators,' (the processes through which change occurs, ie. the characteristics which influence the phenomenon and extent of change, Kazdin and Nock, 2003). Kazdin (2008) pointed out that much evaluation research focuses on whether an intervention works at the expense of understanding how or why it works. Yet in order to maximise therapeutic change, it is necessary to understand both the mechanisms and the variables upon which the effectiveness of those mechanisms depend (*ibid, p3-4*).

The researcher considers this to be a crucial point in the evaluation of the FRIENDS programme and refers the reader back to the realist explanation of outcomes (Chapter Three, Figure 3.1; Robson, 2002). Section 2.5.4 illustrated how the intervention has been associated with success in Australia, the culture for which it was designed, but evidence for its effectiveness outside of this context has been inconsistent, suggesting that moderating factors might play a central role in this discrepancy (Gallegos, 2008). This is supported by two very recent papers: a further Canadian study (Miller et al, 2011) which found no universal intervention effects (N = 253, mean age 9.8), with both the FRIENDS and attention control conditions showing decreased anxiety symptoms; and a review of the existing evidence for the programme (Briesch et al, 2010) which highlights a number of methodological shortcomings alongside the reported successes. Indeed, the latter study identified that there have been no investigations that isolate the different

components of FRIENDS, concluding, "knowledge about the efficacy and necessity of specific aspects of the intervention is limited," (*p162*).

A detailed rationale for assessing mechanisms of change is offered by Kazdin and Nock (2003). They propose that strong associations need to be established between the intervention and the mechanism, and then between the mechanism and therapeutic change. Unfortunately, this study has not afforded the possibility to disaggregate the 'mediating' components and so is unable to identify which (if any) contributed most significantly to the reduction of emotional distress in the IG. The high moderate effect size in emotional distress change scores at best indicates that the FRIENDS 'package' made a contribution comparable with other universal studies (r = .24, Briesch et al, 2010).

Although the programme 'mediators' are unclear, a number of 'moderators' are identified that may have influenced the children's receptiveness to the key components. One explanation for the variability in the results of FRIENDS research may be that the relationship between the proposed mediators and outcomes is moderated by a variable that has not yet been understood (Kazdin and Nock, 2003). Some of the proposed moderators in this study include: the additional reinforcement and modelling of coping skills and relaxation techniques; adaptations to make the language and content more culturally relevant to this sample; the differentiation of activities to suit the learning styles and developmental levels of the children involved; the use of the scrapbook where children had the opportunity to record their responses more freely; the reinforcement of emotional literacy skills through classroom displays; the provision of homework support clubs; individual mentoring to discuss issues arising from the programme; high adult to child ratio during the sessions to talk through the ideas and reinforce learning points; behavioural reinforcement for applying the key skills, and access to support, consultation and reinforcement from the educational psychologist.

As well as maintaining a high degree of treatment integrity, these arrangements reflect a strong emphasis upon interpersonal relationships

between the children, staff and EP. This notion of the 'therapeutic alliance' as a likely mechanism of change has received increasing support in research (Dunsmuir, 2010; DCSF, 2008) and, although beyond the scope of the present study to explore fully, would be a worthwhile component of future FRIENDS evaluations.

Figure 5.1 summarises the components of this analysis for a class of Year 5 children with the demographics described in section 3.6.4:

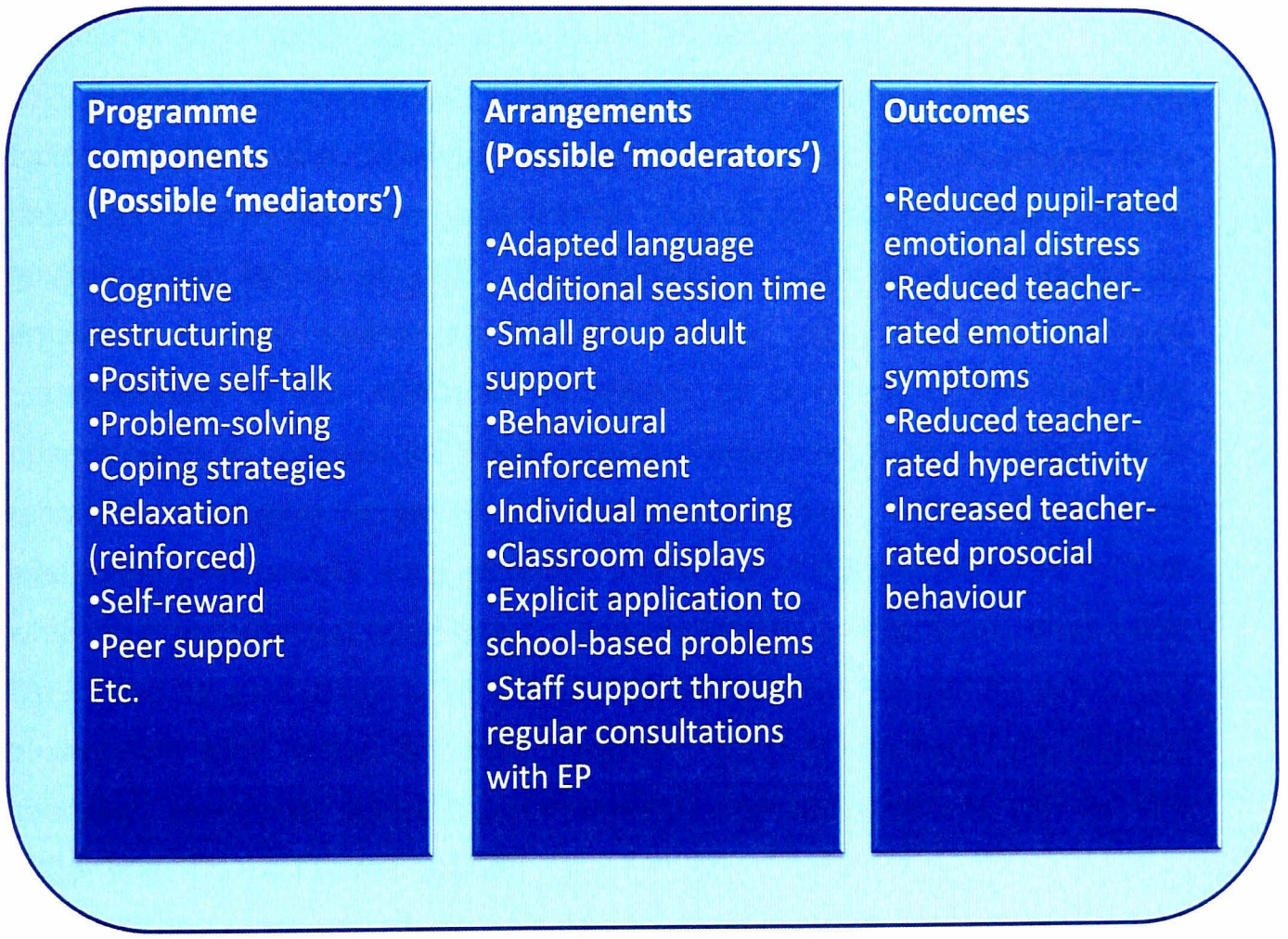


Figure 5.1: Programme components, arrangements and outcomes in this intervention

It is impossible to identify from this analysis any definitive causal directions and as Pawson and Tilley (1997) emphasised, many critical realist interpretations are similarly speculative. However, it would be helpful for future studies to document contextual factors in order to explore patterns in successful moderators.

5.6 Transactional Explanations

A further complication in assessing the direct influence of the intervention components relates to the ripple effects created by the interaction between the project, the children's individual systems and the psychosocial system of the school (Bronfenbrenner, 1979; Miller and Leyden, 1999). Referring back to the SCT model of reciprocal determinism (Bandura, 1986) provides a useful framework to consider these phenomena (Chapter Two, ***Figure 2.1***). While CBT programmes traditionally target the 'internal' personal and behavioural factors of this model, (Stallard, 2010), when applied in a classroom as a universal intervention the programme inevitably impacts upon external environmental factors as well. For example, in addition to the heightened emotional literacy awareness through displays, staff reported enhanced understanding of the children's behaviour as a result of the increased opportunities to discuss individual circumstances, while the children were reportedly displaying increased empathy for each other. It is likely that these interacting phenomena were reflected in the peer relations/prosocial scores, which in turn impacted upon other areas of strength and difficulty, both within the IG and beyond (Biggs et al, 2010). This echoes Kazdin and Nock's (2003) observation that "within a given discipline, multiple pathways (ie. multiple causality, reciprocal causality, bi-directional changes) tend to be the rule rather than the exception" (*p1120*).

It was proposed in section 2.1 that the intervention would boost specific protective factors, which would then interact with individual systems of personal and environmental risk factors (Newman, 2004). The present sample appeared to have a considerable number of such risk factors, when the emotional distress scores, social demographics and educational information are scrutinised, and the results provide some support for the ability of FRIENDS to "buffer adverse experiences" (Frederickson, 2002). Subtle indicators such as the positive change scores across all measures for the IG and the comparatively small number of children moving into or

remaining in 'at risk' ranges on the measures at post-test reinforce this conclusion (Stallard, 2005). Moreover, consultations revealed that several children perceived as having challenging temperaments were able to respond to parts of the programme and demonstrated considerable improvements on the SDQ. This suggests that FRIENDS can be appealing and beneficial to children who reportedly find aspects of school difficult, although case study data would have enabled a more detailed exploration of this supposition (Barlow, Nock and Hersen, 2008). Furthermore, as the CG also showed improvements (especially in conduct and prosocial scores), this supports the notion of a complex system of interactions (Bronfenbrenner, 1979). Indeed, the researcher's philosophy evolved as it progressed, from a distanced, post-positivist evaluation to the more participatory, 'action research' stance advocated by complexity theory (Morrison, 2002; Cohen et al, 2007), although the methodology did not equally evolve to capture this intricacy.

5.7 Does This Study Support the Use of CBT With Primary-Aged Children?

In section 2.4, some of the issues surrounding the application of CBT with children were raised, including whether they possess the cognitive capacity to engage with the strategies taught (Grave and Blissett, 2004). From the researcher's anecdotal direct observations of the children, it appeared that the Year 5 pupils in particular were able to engage effectively with the programme components although it is not possible to report how well they generalised skills, such as cognitive restructuring, outside of the FRIENDS sessions. Indeed, the question of whether CBT programmes lead to genuine cognitive change or merely reflect the outcomes of behavioural reinforcement has been the subject of much debate (Durlak et al, 1991; Stallard, 2002). Some informal anecdotal evidence that suggested individuals were beginning to internalise both the cognitive and behavioural strategies was provided by a girl in the IG; on witnessing her teacher becoming frustrated with the computer she spontaneously remarked:

"Mrs W, you need to relax and use your green thoughts!"

However, it was the researcher's observation at the end that most of the children were still applying general rather than specific 'green thoughts' to given scenarios, which suggested that this aspect would require ongoing reinforcement.

The researcher's observations and consultations with staff indicated that the younger Year 3s had greater difficulty with accessing the emotional vocabulary and required considerable reinforcement to make the distinctions between thoughts, feelings and behaviours, reflecting Stallard's (2005) observation that some children may find it difficult to engage with the cognitive component. This was particularly an issue for the children with English as an additional language whose Cognitive Academic Language Proficiency was at an early stage (Cummins, 1984). This is perhaps reflected in the fact that no perceptible movement was observed in the self-reports in the pilot study (although this may equally have been a problem of accessing the measures reliably; Dockrell, Lewis and Lindsay, 2000). In both year groups, the programme needed appropriate cultural and developmental differentiation (as documented previously) for the children to access the ideas (Grave and Blissett, 2004), but the Year 5s were able to proceed more independently once the strategies had been modelled to them. Another factor that appeared to facilitate this was the external reinforcement of the internal 'therapeutic' techniques through peer support (Barrett, 2004). In summary, Zimmer-Gembeck and Skinner's (2011)'s observation is supported by the pattern of results in this study:

"It is likely that [in middle childhood] children's growing abilities to take the perspective of others allows them to better co-ordinate their coping with social partners and to more effectively negotiate interpersonal stressors." (p12).

5.8 Is FRIENDS Effective as a Universal Intervention?

Chapter Two highlighted some of the reservations that have arisen regarding the universal application of FRIENDS (Rose, Miller and Martinez, 2009). One of the difficulties with assessing change within universal groups is that floor effects for children with initial low anxiety are likely to confound the strength of impact on those with elevated levels (Briesch et al, 2010). With a larger sample size, it may have been illuminating to separate the data for these different groups of children (see Gallegos, 2008) and to compare effects between them. However, the positive results across all measures for the IG, together with the data on preventive effects, provide some support for the effectiveness of the programme – or at least, suggest that, in interaction with other psychosocial phenomena, it "did more good than harm" (Hammersley, 2005, *p.85*).

These results are important because the demographic and educational data for these children, together with their 'elevated' levels of emotional distress, suggest that they might be one of the vulnerable groups identified in section 1.1. (DCSF, 2008), and it has been estimated that around 40% of children with a psychological difficulty are not receiving specialist treatment (Rait et al, 2010). For this reason, it is suggested that with appropriate training, school staff can play a vital role in identifying and supporting such children through their daily interactions and possibly in pre-empting difficulties before they intensify (Briesch et al, 2010). Furthermore this study supports a double advantage of universal over targeted implementation: not only does it avoid the stigma associated with discussing anxiety and mental health problems, (Frederickson et al, 2009), but it offers children the opportunity to support each other with the reinforcement of skills (Barrett, 2004).

5.9 The Role of the EP in Supporting Mental Health Initiatives

In Chapter Two it was proposed that the EP's unique working knowledge of school systems and how these impact upon behaviour places them in a strong position to support staff with the delivery of non-selective CBT interventions at a preventive level (Rait et al, 2010). It is suggested that this is a cost-efficient way of maximising EP time and input, enabling more children to be supported and possibly reducing the demand on SEN resources (Squires, 2001). As growing evidence supports the effectiveness of CBT with psychological difficulties such as anxiety and depression "at the milder end of the spectrum," Rait et al (2010) suggest that EPs could play a "distinct supervision role" in this context (*p117*).

The researcher endorses the above view that EPs *should* be part of a consultation structure around the implementation of programmes such as FRIENDS. Although Barrett et al (2001) found no difference in the effectiveness between teacher and psychologist delivery, more recent evidence suggests that the mean effect size for trained practitioners as implementers (ES = .56) is twice that for teachers or school staff alone (Briesh et al, 2010). Not only may EPs bring an embedded understanding of the psychological theories underpinning CBT, but they are well-placed to use supportive networks to ensure effective and ethical delivery (Squires, 2010). This was salient in the present study following the heightened levels of emotional distress reported on the PI-ED which required careful interpretation and monitoring, involving a consultative network between the researcher, her supervisor, school staff and parents, where necessary.

A further role for the EP concerns thinking around how the CBT model can inform organisational practice, including helping adults to manage their own emotional reactions when working in challenging school contexts (Rait et al, 2010). Through careful relationship building, the EP may become alert to how psychosocial factors within the school might be operating to impact upon both staff and pupil mental health, and how both might be supported accordingly

(DCSF, 2009). Thus the focus would be less upon individual intervention and more upon how non-selective CBT programmes like FRIENDS can support *groups* of children, taking into account their *collective* risk and protective factors (Rait et al, 2010). SCT or Miller and Leyden's (1999) coherent framework provide useful models to guide observations (***Figure 5.2***).

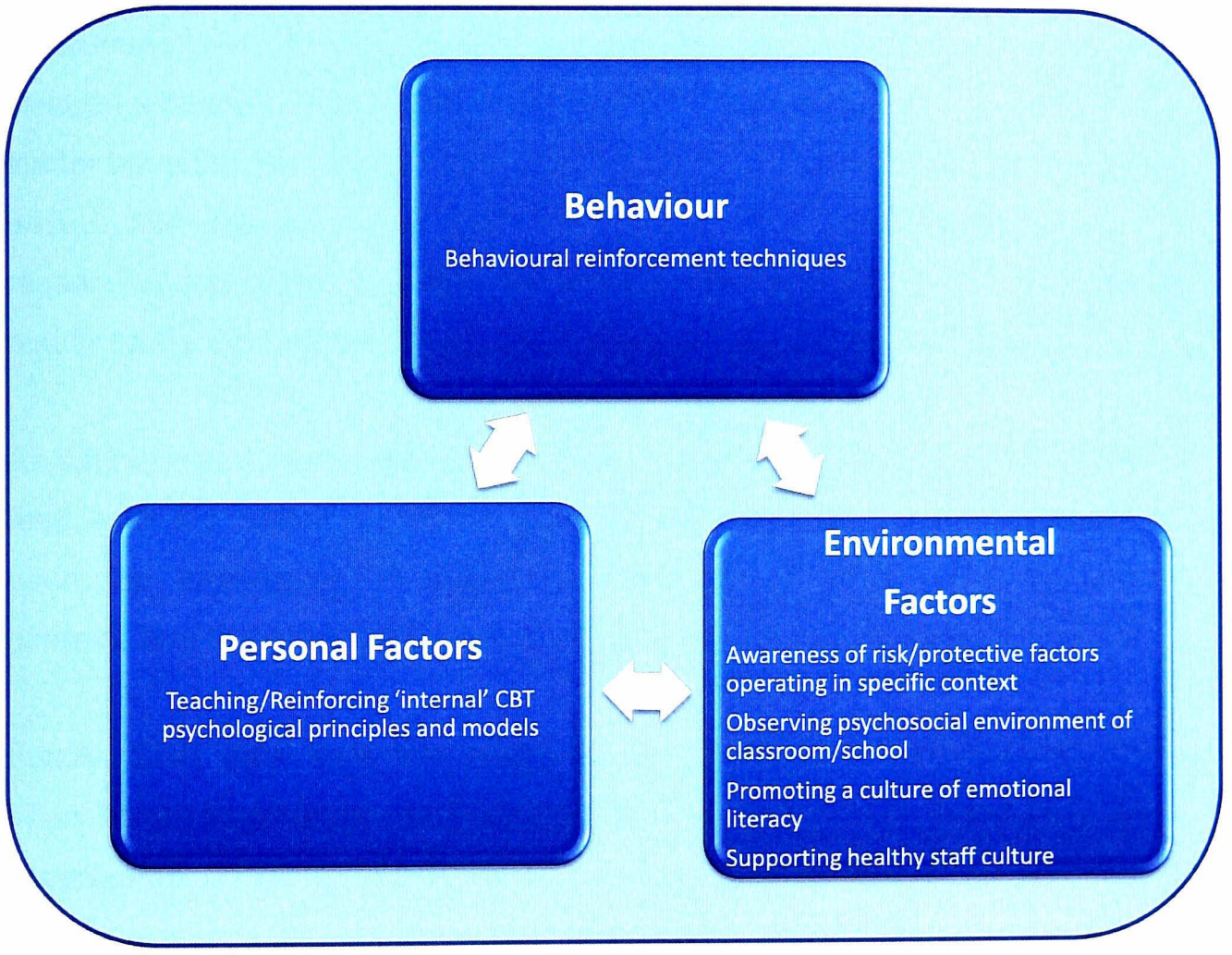


Figure 5.2: An example of how the EP might support the delivery of a CBT intervention in schools.

5.10 Stakeholders

Informal feedback from the project school has been overwhelmingly positive about the FRIENDS programme and participating staff are keen to disseminate training across the school so that techniques can be universally reinforced. The IG will receive their booster sessions, the CG is currently receiving the programme and the PI-ED will be re-administered at the end of the summer term to evaluate its impact upon this group and to identify any children who may require further intervention. The researcher and the group leader have planned a summative consultation to address these issues and to ensure that the school is appropriately supported with next steps. The researcher also plans to revisit the school in the autumn term to present the results to the staff and children, and a summary will be sent to parents.

As part of the current countywide rollout of TaMHS, the researcher plans to feed back the outcomes of this study to the EPS, with particular emphasis upon the importance of assessing treatment integrity, individual school contexts and the potential factors that might moderate programme outcomes.

Finally, SDQ data from this study will be aggregated with that from other educational psychology doctoral research to inform the national D & R Programme.

5.11 Summary of Methodological Strengths and Limitations

Strengths

The present study has addressed a key problem identified in previous UK studies of FRIENDS by including a control group in its design (Stallard, 2008). This has permitted a fuller exploration of whether changes in emotional distress and aspects of teacher-reported behaviour may be attributed to the intervention or to confounding variables. By preserving intact classes, the study demonstrates high ecological validity by investigating the effects of

FRIENDS on children in their natural classroom environment (Robson, 2002). This increased the chance that social phenomena would unfold more authentically (Mertens, 2010), which may be detected in the diminished peer problems scores at post-test, for example. Furthermore, the researcher's immersion in the programme implementation enabled first-hand observation of the children's responses and permitted both scrutiny and manipulation of the moderators influencing this process (Pawson and Tilley, 1997). The rich information that was gained from this approach challenges the notion of the distanced post-positivist, RCT protocol as the 'gold standard' in education research and highlights the benefits of evaluating processes as well as outcomes (Petticrew and Roberts, 2003).

The presence of the researcher also provided consultative supervision and helped to establish a trusted alliance between EP, school staff and pupils. This was considered to be an important factor in fulfilling ethical requirements (See section 3.14). Meanwhile, aspects of the methodology have reflected some robust evaluative procedures (Gersten et al, 2005), including: high treatment integrity; minimising diffusion of treatments; attempts to triangulate findings through self and teacher-report; and matching the unit of statistical analysis to the research question. Data analysis techniques also took into account the strengths and limitations of methodology and design in the adoption of the change score procedure.

Limitations

Although the absence of random allocation may have advantages in preserving the children's natural interactions, it introduced selection biases that were difficult to control for in the statistical analysis and rendered it difficult to exclude validity threats such as history and maturation (Robson, 2002). Moreover, the lack of blinding introduced a confounding variable that severely restricted the ability to attribute any reported changes to the intervention (Cohen et al, 2007). This is compounded by the limitations associated with self and proxy reports, as well as the difficulties with operationalising ambiguous constructs such as anxiety, self-concept and self-

efficacy (Dockrell et al, 2000; Wigelsworth et al, 2010). Problems associated with the construct validity of the PI-ED in particular have been raised; even though the outcomes reflect those of other anxiety measures used in FRIENDS research and are strongly correlated with self-perception change scores as would be predicted by self-efficacy theory (Bandura, 1986). The two major limitations of this study are therefore that the researcher cannot be confident that any of the observed changes were attributable to the intervention, (although the pattern of results indicates some promising trends), and that it has not been possible to identify with more certainty the precise mechanisms that have been operating (even though a number of likely moderators have been highlighted).

This study has also focused upon a very small sample, which may not be representative of the wider population, and this affects the ability to generalise conclusions (Cohen et al, 2007). However, trends suggest that evaluative practice may move away from the emphasis upon generalisability, towards more local applied research that focuses on the development of specific emotional difficulties in context, (Rait et al, 2010; DfE, 2011).

Another limitation involves the lack of parental involvement or measures which has restricted analysis of whether changes were generalised to home contexts, and omitted the parental component of the programme which has been identified as requiring further research (Briesch et al, 2010).

Finally, this study requires further long-term follow-up to establish whether improvements are maintained.

Future Directions

Table 5.1 indicates some questions arising out of the present study and some possible methods to address them:

Research Question	Possible Methods
1. Which, if any, of the FRIENDS components are most effective in reducing emotional distress?	• Measures of cognitive restructuring, problem-solving, coping and relaxation skills taken at several points throughout the programme (Kazdin and Nock, 2003). • Varying the emphasis on different programme components • Mixed methods to gain children's/staff/parental views on the most helpful components. • Qualitative analysis of the 'therapeutic relationship'
2. Do selected moderating influences affect programme outcomes?	• Quasi-experimental design with one class receiving only FRIENDS and the other, FRIENDS plus potential moderators (see section 5.5). • Replication of small scale studies in other contexts to explore patterns in successful moderators. • Using additional comparison groups
3. Does FRIENDS improve academic self-perceptions in specific domains?	• Single case experimental design evaluating domain-specific self-perceptions. • Mixed methods to explore application of the components in children's particular areas of difficulty and to investigate the "meanings" that individuals attach to their performance (Usher and Pajares, 2008).
4. Does FRIENDS have long-term educational benefits?	• Scrutiny of attainment data over time for intervention/control/comparison groups.
5. Are programme benefits maintained over time?	• T3 standardised measures for both the IG and CG, including 12 months+ data.

Table 5.1 Topics for future research and possible methods to address them

6 CHAPTER SIX: CONCLUSION

This study set out to evaluate the impact of a universal FRIENDS intervention upon the reduction and prevention of emotional distress, the enhancement of academic self-perceptions and teacher-rated pupil behaviour. The aims were to contribute to the growing body of knowledge about universal mental health prevention programmes; to provide evaluation data for the local authority; to supply data for the National D & R Programme and to provide feedback to the project school about the programme's effectiveness within their specific context. Outcomes indicated that the class of Year 5 children participating in a universal FRIENDS intervention displayed significantly lower levels of self-reported emotional distress than a parallel Year 5 wait-list control group. There were no significant changes within or between groups for self-reported academic self-perceptions. Both classes showed significantly improved teacher-reported behaviour, with the IG showing significantly improved hyperactivity scores in relation to the CG. There were also significant improvements for both groups in prosocial behaviour. Finally, the number of children in the IG moving out of the 'range of concern' for both emotional distress and academic self-perceptions indicates a preventive effect for the programme or at least suggests that "more good than harm" resulted from its application (Hammersley, 2005).

A Unique Contribution

Within the limitations of its protocol, this study supports the previous UK research on FRIENDS by indicating that an intervention group's reduction in emotional distress was significantly greater than a control group's, lending some credence to the effectiveness of the intervention in this context. The study has also highlighted trends in relation to hyperactivity and prosocial behaviour, which may be interesting avenues for future research. However, the perils of inferring causal inferences from quasi-experimental designs have

been thoroughly acknowledged and these results should therefore be interpreted with caution (Stevens, 1999).

The study has also highlighted the importance of exploring processes as well as outcomes when evaluating therapeutic programmes (Kazdin, 2007). It has been suggested that particular components of FRIENDS or aspects of its contextual implementation may be critical in producing beneficial effects, but thus far, research has been unable to specify these (Briesh et al, 2010). The present study has highlighted some potential 'moderators' that may be helpful to explore in future studies.

The lack of evidence for the effects of CBT upon additional cognitive variables has been highlighted as a gap in previous research (Grave and Blissett, 2004) and this study attempted to address this in the measurement of academic self-perceptions. However, the difficulty of operationalising abstract constructs such as self-concept and self-efficacy has been emphasised and the researcher has drawn attention to the need to differentiate these constructs more specifically, possibly using additional mixed methods to explore them more reliably.

The study provides important information for the stakeholders involved. The outcomes contribute to the growing evidence for the application of universal therapeutic interventions in schools (D&R programme; Adi et al, 2007). As the intervention showed some effectiveness with a population that may be described as vulnerable, this is likely to support the local authority's promotion of FRIENDS in schools with similar catchments. Meanwhile, the project school's positive experience has resulted in a firm commitment to establishing the FRIENDS principles within their school PSHE curriculum. The challenge will now be to involve parents further in the programme delivery and evaluation.

Finally, with its acknowledgement of the importance of environmental as well as cognitive components, the researcher has suggested a model for how EPs might support the application of CBT in schools.

In conclusion, the links between emotional distress and academic achievement that inspired this study continue to be echoed in contemporary governmental literature (DfE, 2011). Through this project, the researcher has come to appreciate the full complexity of investigating these phenomena, including how to operationalise intangible constructs and establishing continuity between purpose, research questions, epistemology and design. The limitations of the current protocol in evaluating emotional distress and academic self-perceptions effectively have been recognized, and the author aspires to the challenge of embracing a wider range of methodologies in her future work as an applied research practitioner.

7 REFERENCES

REFERENCES

Adi, Y., Killoran, A., Janmohamed, K, & Stewart-Brown, S. (2007). *Systematic review of the effectiveness of interventions to promote mental wellbeing in children in primary education. Report1: universal approaches (non-violence-related outcomes).* London: National Institute for Health and Clinical Excellence.

Allen, J.L., & Rapee, R.M. (2005). Anxiety Disorders in P. Graham, (Ed), *Cognitive Behaviour Therapy for Children and Families (Second Edition).* Cambridge University Press.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV (4th edition).* Washington DC: American Psychiatric Association.

Amstadter, A.B. Emotion regulation and anxiety disorders. *Journal of Anxiety Disorders 22 (2)*, 211-221.

Baldwin, J.S. & Dadds, M.R. (2008). Examining alternative explanations of the covariation of ADHD and anxiety symptoms in children: a community study. *Journal of Abnormal Child Psychology 36*, 67-79.

Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychology Review* , 84 (2), 191-215.

Bandura, A. (1986). *Social foundations for thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1988). Self-efficacy conception of anxiety. *Anxiety Research*, 1, 77-98.

Bandura, A. (1997). *Self-efficacy: The Exercise of Control.* Freeman. New York.

Barlow, D.H., Nock, M.K., & Hersen, M. (2008*). Single Case Experimental Designs: Strategies for Studying Behaviour Change (3rd edition*). Boston, MA: Allyn and Bacon.

Barrett, P. (2000). Treatment of Childhood Anxiety: Developmental Aspects. *Clinical Psychology Review 20 (4)*, 479-494.

Barrett, P. (2004). *FRIENDS for life: Group leader's manual for children.* Brisbane, Australia: Australian Academic Press.

Barrett, P.M. & Pahl, K.M. (2006). School-based Intervention: Examining a Universal Approach to Anxiety Management. *Australian Journal of Guidance*

and Counselling: Special Issue: Mindmatters Plus: A national mental health initiative in secondary schools 16(1), 55-75.

Barrett, P.M. & Turner, C.M. (2001). Prevention of anxiety symptoms in primary school children: Preliminary results from a universal school-based trial. British *Journal of Clinical* Psychology 40, 399-410.

Barrett, P.M., Farrell, L.J., Ollendick, T.H., & Dadds, M. (2006). Long term outcomes of an Australian universal prevention trial of anxiety and depression symptoms in children and youth: An evaluation of the Friends program. *Journal of Clinical Child and Adolescent Psychology 35(3),* 403-411.

Barrett, P.M., Lock, S. & Farrell, L. (2005). Developmental differences in universal preventive intervention for child anxiety. *Clinical Child Psychology and Psychiatry 10(4)*, 539-555.

Barrett, P.M., Lowry-Webster, H.M., & Turner, C.M. (1999). *FRIENDS program social validity and treatment integrity scales*. Brisbane: Australian Academic Press.

Beck, A., Emery, G. & Greenberg, R. (1985). *Anxiety Disorders and Phobias*. New York: Guilford.

Becker, L.A. (1999). *Effect size calculators.* http://www.uccs.edu/~faculty/lbecker/ Retrieved on 15.03.11.

Beesdo, K., Knapp, S. & Pine, D.S. (2009). Anxiety and anxiety disorders in children and adolescents: developmental issues and implications for DSM-V. *Psychiatric Clinics of North America 32 (3)*, 483 – 524.

Begley, A. (2000). The educational self-perceptions of children with Down Syndrome. In A. Lewis & G. Lindsay. *Researching Children's Perspectives.* Buckingham: Open University Press.

Bernstein, G.A., Layne, A.E., Egan, E.A & Tennison, D.M. (2005). School-based interventions for anxious children. *Journal of American Academy of Child and Adolescent Psychiatry, 44 (11),* 1118-1127.

Bevans, K.B. & Forrest, C.B. (2009). The reliability and validity of children and adolescents' self-reported health. *Economic Evaluation in Child Health* (Oct), 33-55.

Biggs, Nelson and Sampilo (2010). Peer-relations in the anxiety-depression link: test of a mediation model. *Anxiety, Stress and Coping: An International Journal 23 (4),* 431-447.

Bolton, D. (2005). Cognitive behaviour therapy for children and adolescents: some theoretical and developmental issues. In P. Graham (Ed), *Cognitive*

Behaviour Therapy for Children and Families: Second Edition. Cambridge University Press.

Bong, M. & Skaalvik, E.M. (2003). Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educational Psychology Review, 15 (1)*, 1-40.

BPS, Division of Education and Child Psychology (2002). *Framework for psychological assessment and intervention. British Journal of Psychology, Professional Practice Guidelines.* Leicester: BPS.

Brace, N., Kemp, R. & Snelgar, R. (2006). *SPSS For Psychologists (3rd edition)*. Hampshire, UK: Palgrave Macmillan.

Briesch, A.M., Hagermoser Sanetti, L.M & Briesch, J.M (2010). Reducing the prevalence of anxiety in children and adolescents: An evaluation of the evidence base for the FRIENDS for Life Program. *School Mental Health 2,* 155-165.

British Psychological Society (2006). *Code of Ethics and Conduct*. Leicester: BPS.

Bronfenbrenner, U. (1979). *The ecology of human development.* Cambridge, MA: Harvard University Press.

Brunner, M., Keller, U., Dierendonck, C., Reichert, M. Ugen, S., Fischbach, A. & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology, 102 (4),* 964-981.

Burden, R. (1998a). Assessing Children's Perceptions of Themselves as Learners and Problem-solvers: The Construction of the Myself-As-Learner Scale (MALS). *School Psychology International, 19 (4),* 291-305.

Burden, R. (1998b). Myself As Learner Scale (MALS). In N. Frederickson and R.J. (Seàn) Cameron *Psychology in Education Portfolio*. London: NFER-Nelson.

Burden, R. (1999). Children's Self-Perceptions: Introduction. In N. Frederickson and R.J. (Seàn) Cameron *Psychology in Education Portfolio.* London: NFER-Nelson

Butler, A.C., Chapman, J.E., Forman, E.M & Beck, A.T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review 26 (1)*, 17-31.

Caprara, G.V., Vecchione, M., Alessandri, G., Gerbino, M. & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology, 81(1)*, 78-96.

Carr, A. (2000). *What works with children and adolescents? A critical review of psychological interventions with children, adolescents and their families.* Routledge.

Carr, A. (2006). *The Handbook of Child and Adolescent Clinical Psychology: A Contextual Approach* (2nd Edition). Routledge. London, New York.

Cartwright-Hatton, S., Roberts, C., Chitsabesan, P, Fothergill, C. & Harrington, R. (2004). Systematic review of the efficacy of cognitive behaviour therapies for childhood and adolescent disorders. *British Journal of Clinical Psychology 43,* 421-436.

Chalmers, I. (2003). Trying to do more good than harm in policy and practice: the role of rigorous, transparent, up-to-date evaluations. *Annals of the American Academy of Political and Social Science 589*, 22-40.

Chatterji, M. (2008). Synthesizing Evidence from Impact Evaluations in Education to inform action. *Educational Researcher 37 (1),* 23-26.

Choi, N. (2005). Self-efficacy and self-concept as predictors of college students' academic performance. *Psychology in the Schools 42 (2),* 197-205.

Clark, L.A. & and Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Child Psychology 100,* 316-336.

Cohen, L., Manion, L. & Morrison, K. (2007). *Research Methods in Education (6th edition).* London: Routledge.

Compton, S.N., March, J.S., Brent, D., Albano, A., Weersing, R. & Curry, J. (2004). Cognitive-Behavioral Psychotherapy for Anxiety and depressive Disorders in Children and Adolescents: An Evidence-Based Medicine Review. *Journal of the American Academy of Child and Adolescent Psychiatry, 43 (8),* 930-959.

Connell, J., Barkham, M., Stiles, W.B., Twigg, E., Singleton, N., Evans, O. & Miles, J.N. (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points and comparison with the CIS-R. *British Journal of Psychiatry 190*, 69-74.

Conover W.J., (1999). *Practical Nonparametric Statistics* (3rd edition). Wiley.

Cook, T.D. & Campbell, D.T. (1979*). Quasi-Experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.

Coolican, H. (2009). *Research Methods and Statistics in Psychology.* London: Hodder Education.

Costello, E.J., Mustillo, S., Erkanli, A., Keeler, G. & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry 60*, 837-844.

Craven, R., Marsh, H. & Burnett, P.C. (2004). Breaking the self-concept enhancement conundrum: re-conceptualising the next generation of self-concept enhancement research in *NZARE/AARE Conference 2003: Educational Research Risks and Dilemmas, 29 November – 3 December* Auckland, New Zealand.

Craven, R.G. & Marsh, H.W. (2008). The centrality of the self-concept construct for psychological well-being and unlocking human potential: Implications for child and educational psychologists. *Educational and Child Psychology 25 (2)*, 104-118.

Cummins, J. (1984). *Bilingualism and Special Education: Issues in assessment and pedagogy*. Clevedon: Multilingual Matters.

Dancey, C.P & Reidy, J. *Statistics Without Maths for Psychology (4th edition)*. Harlow: Pearson.

Davidson, J. (2008). *Children and young people in mind: the final report of the National CAMHS review*. London: DCSF/DOH.

DCSF (2005). *Social and Emotional Aspects of Learning: improving behaviour, improving learning*. DCSF Publications.

DCSF (2008). *Improving the mental health and psychological well-being of children and young people: National CAMHS Review interim report.* London: DCSF.

DCSF (2009). *Learning From Targeted Mental Health in Schools Phase 1 Pathfinders. Summary Report*. Nottingham: DCSF.

DCSF (2010). *Guidance on commissioning targeted mental health and emotional wellbeing services in schools*. Nottingham: DCSF Publications.

Degnan, K.A., Almas, A.N & and Fox, N.A. (2010). Temperament and the environment in the etiology of childhood anxiety. *Journal of Child Psychology and Psychiatry 51 (4)*, 497-517.

DfE (2011). Support and aspiration: A new approach to special educational needs and disability – a consultation. London: DfEE Publications

DfES (2001). *Promoting children's mental health in early years and school settings.* Nottingham: DfEE Publications.

DfES (2004). *Every Child Matters: Change for Children*. HM Government.

Dockrell, J., Lewis, A & Lindsay, G. (2000). Researching Children's Perspectives: a psychological dimension. In A. Lewis & G. Lindsay. *Researching Children's Perspectives*. Buckingham: Open University Press.

DOH (2004). *National Service Framework for Children, Young People and Maternity Services*. London: DH Publications.

Dryden, W & David, D. (2008). Rational Emotive Behaviour Therapy: Current Status. *Journal of Cognitive Psychotherapy 22 (3),* 195-209.

Duchesne, S., Vitaro, F., Larose, S. & Tremblay, R.E. (2008). Trajectories of Anxiety during Elementary-School Years and the Prediction of High School Noncompletion. *Journal of Youth and Adolescence, 37 (9),* 1134-1146.

Dunsmuir, S. (2010). *Cognitive Behavioural Therapy for Children in schools.* "Leading Edge Psychology" presentation given at University College London on Wednesday 8th September 2010.

Durlak, J.A., Fuhrman, T., & Lampman, C. (1991). Effectiveness of cognitive-behavior therapy for maladapting children: A meta-analysis. *Psychological Bulletin*, 110, 204-214.

Ellis, A. (1980). Rational-Emotive Therapy and Cognitive Behaviour Therapy: Similarities and Differences. *Cognitive Therapy and Research, 4 (4)*, 325 - 340.

Farrell, P., Woods, K., Lewis, S., Rooney, S., Squires, G. & O'Connor, M. (2006). *A Review of the functions and contribution of educational psychologists in England and Wales in light of "Every Child Matters: Change for Children."* London: HMSO.

Fergusson, D.M. & Lynskey, M.T. (1996). Adolescent resiliency to family adversity. *Journal of Adolescent Child Psychology and Psychiatry, 37,* 281-292.

Fitzmaurice, G. (2001). A conundrum in the analysis of change. *Nutrition 17 (4),* 360-361.

Fitzmaurice, G.M., Laird, N.M & Ware, J.H. (2004). *Applied longitudinal analysis.* Hoboken, NJ: Wiley.

Flavell, J.H., Flavell, E.R., & Green, F.L. (2001). Development of children's understanding of connections between thinking and feeling. *Psychological Science, 12,* 430-432.

Frederickson, N. (2002). Evidence-based practice and educational psychology. *Educational and Child Psychology 19 (3)*, 96-111.

Frederickson, N., Dunsmuir, S. & Baxter, J. (2009). *Measures of Children's Mental Health and Psychological Wellbeing: A Portfolio for Education and Health Professionals: Introduction*. London: GL Assessment.

Fröjd, S.A., Nissenen, E.S., Pelkonen, M.U.I., Marttunen, M.J., Koivisto, A. & Kaltiala-Heino, R. (2008). Depression and school performance in middle adolescent boys and girls. *Journal of Adolescence 31 (4)* 485-498.

Frydenberg, E. (2008). *Adolescent Coping: Advances in theory, research and practice.* London: Routledge.

Gallegos, J. (2008). *Preventing Childhood Anxiety and Depression: Testing the Effectiveness of a School-based program in Mexico*. (Dissertation for Doctor of Philosophy). University of Texas at Austin. http://gradworks.umi.com/33/41/3341564.html Retrieved 09.07.10.

Gay, L.R., Mills, G.E., Airasian, P.W. (2009). *Educational Research: Competencies for analysis and application (9th edition).* Upper Saddle River, NJ: Merrill Prentice Hall.

Glass, G.V., Peckham, P.D. & Sanders, J.R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance. *Review of Educational Research 42,* 237-288.

Gliner, J.A., Morgan, G.A. & Harmon, R.J., (2003). Pre-test-Posttest Comparison Group Designs: Analysis and Interpretation. *Journal of the American Academy of Child and Adolescent Psychiatry, 42 (4),* 500-503.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. Journal of Child Psychology and Psychiatry, 38 (581-586).
Graham, P. (2005). *Cognitive Behaviour Therapy for Children and Families (Second Edition)*. Cambridge University Press.

Graham, P. (2005). *Cognitive Behaviour Therapy for Children and Families (2nd edition)*. Cambridge: Cambridge University Press.

Grave, J. & Blissett, J. (2004). Is cognitive behaviour therapy developmentally appropriate for young children? A critical review of the evidence. *Clinical Psychology Review*, 24, 399-420.

Green, H., McGinnity, A., Meltzer, H., Ford, T & Goodman, R. (2004). Mental health of children and young people in Great Britain, 2004. A survey carried out by the Office for National Statistics. London: HMSO

Hacker, D.J., Dunlosky, J. & Graesser, A.C. (2009). A Growing Sense of "Agency" in D.J. Hacker & A.C. Graesser (eds) *Handbook of metacognition in education*. New York: Routledge.

Hammerness, P., Gellere, D., Petty, C., Lamb, A., Bristol, E. & Biederman, J. (2010). Does ADHD moderate the manifestation of anxiety disorders in children? *European Child and Adolescent Psychiatry 19 (2)*, 107-112.

Hammersley, M. (2005). Is the evidence-based practice movement doing more good than harm? Reflections on Iain Chalmers' case for research-based policy making and practice. *Evidence and Policy 1 (1),* 85-100.

Hampel, P., & Petermann, F. (2006). Perceived stress, coping and adjustment in adolescents. *Journal of Adolescent Health 38*, 409-415.

Hansford, B.C. & Hattie, J.A. (1982). The Relationship between self and Achievement/Performance Measures. *Review of Educational Research 52,* 123-142.

Harwell, M.R. & Gatti, G.G. (2001). Rescaling Ordinal Data to Interval Data in Educational Research. *Review of Educational Research 71,* 105-131

Hawes, D.J & Dadds, M.R. (2004). Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. *Australian and New Zealand Journal of Psychiatry, 38,* 644-651.

Health Professions Council, (2008). *Standards of conduct, performance and ethics*. London: HPC.

Howitt, D. & Cramer, D. (2011*). Introduction to Statistics in Psychology (5th edition)*. Harlow: Pearson

Huber, B.R. Reporting statistics in psychology. http://huberb.people.cofc.edu/Guide/Reporting_Statistics%20in%20Psychology.pdfs. Retrieved on 01.03.11.

Hughes, J.N. (2000). The essential role of theory in the science of treating children: beyond empirically supported treatments. *Journal of School Psychology 38 (4),* 301-330.

James, A., Soler, A., & Weatherall, R. (2005). Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database of Systematic Reviews 2005, Issue 4.* Wiley and Sons Ltd.

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education, 38*, 1217-1218.

Kantowitz, B.H., Roediger, H.L. Elmes, D.G. (2009). *Experimental Psychology (9th edition).* Belmont, CA: Wadsworth, Cengage Learning.

Kazdin, A.E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology 3,* 1-27.

Kazdin, A.E. & Nock, M.K. (2003). Delineating mechanisms of change in child and adolescent therapy: methodological issues and research recommendations. *Journal of Child Psychology and Psychiatry 44(8),* 1116-1129.

Kendall, P.C. & Choudhury, M.S. (2003). Children and adolescents in cognitive-behavioral therapy: Some past efforts and current advances, and the challenges in our future. *Cognitive Therapy and Research 27*, 89-104.

King, N.J., Heyne, D. & Ollendick, T.H. (2005). Cognitive-Behavioral Treatments for Anxiety and Phobic Disorders in Children and Adolescents: A Review. *Behavioral Disorders*, 30 (3), 241-257.

Koskelainen, M. Sourander, A & Kaljonen, A. (2001). The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. *European Child and Adolescent Psychiatry, 9*, 277-284.

Kraaij, V., Garnefski, N., de Wilde, E.J., Dijkstra, A., Gebhardt, W., Maes, S. & ter Doest, L. (2003). Negative life events and depressive symptoms in late adolescence: bonding and cognitive coping as vulnerability factors. *Journal of Youth and Adolescence 32 (3)*, 185-193.

Kurtz, (2004). *What works in promoting children's mental health: The evidence and implictaions for Sure Start local programmes*. London: DfES.

Langley, A.K., Bergman, R.L., McCracken, J., & Piacentini, J.C. (2004). Impairment in childhood anxiety disorders: Preliminary examination of the Child Anxiety Impact Scale – Parent version. *Journal of Child and Adolescent Psychopharmacology*, 14, 105-114.

Last, C.G., Hansen, C. & Franco, N. (1997). Anxious children in adulthood: A prospective study of adjustment. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36, 645-652.

Lau, J.Y.F., Eley, T.C. & Stevenson, J. (2006). Examining the state-trait anxiety relationship: a behavioral genetic approach. *Journal of Abnormal Child Psychology 34 (1),* 19-27.

Lewis, A. & Lindsay, G. (2000). *Researching Children's Perspectives.* Buckingham: Open University Press.

Liddle, I. & Macmillan, S. (2010). Evaluating the FRIENDS programme in a Scottish setting. *Educational Psychology in Practice* 26 (1), 53-67.

Liew, J., McTigue, E.M., Barrois, L. & Hughes, J.N. (2008). Adaptive and effortful control and academic self-efficacy beliefs on achievement: A longitudinal study of 1st through 3rd graders. *Early Childhood Research Quarterly, 23 (4)*, 525-526.

Limbers, C.A., Newman, D.A. & Varni, J.W. (2008). Factorial Invariance of Child Self-report across age subgroups: A confirmatory factor analysis of ages 5-16 years utilizing the PedsQL 4.0 Generic Core Scales. *Value In Health 11 (4)*, 659-668.

Lock, S. & Barrett, P.M. (2003). A longitudinal study of developmental differences in universal preventive intervention for child anxiety. *Behaviour Change, 20 (4),* 183-199.

Lowry-Webster, H., Barrett, P.M. & Dadds, M.R. (2001). A universal prevention trial of anxiety and depressive symptomatology in childhood: preliminary data from an Australian study. *Behaviour Change*, 18 (1), 36-50.

Lowry-Webster, H., Barrett, P.M. & Lock, S. (2003). A universal prevention trial of anxiety symptomatology during childhood: results at one year follow-up. *Behaviour Change*, 20 (1), 25-43.

March, J.S. (1997). Manual for the Multidimensional Anxiety Scale for Children (MASC). Toronto: Multi-Health Systems.

Mark, M.M. (2010). Emergence in and from quasi-experimental design and analysis. In S.N. Hesse-Biber and P. Leavy (eds). *Handbook of Emergent Methods.* New York: Guilford Press.

Marsh, H.W. (1990). The Structure of Academic Self-Concept: The Marsh/Shavelson Model. *Journal of Educational Psychology 82 (4),* 623-636.

Marsh, H.W. (2008). A Multidimensional, Hierarchical Model of Self-Concept: An important facet of personality in G.J. Boyle, G. Matthews & D.H. Saklofske (eds). *The SAGE Handbook of Personality Theory and Assessment Vol 1: Personality Theories and Models.* London: SAGE.

Marsh, H.W. & Martin, A.J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81 (1),* 59-77.

Marsh, H.W. & Yeung, A.S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology, 89,* 41-54.

Martinovich, Z., Saunders, S. & Howard, K.L. (1996). Some comments on 'Assessing clinical significance.' *Psychotherapy Research 6*, 124-132.

Massetti, G.M., Lahey, B.B., Pelham, W.E., Loney, J., Ehrhardt, A., Lee, S.S., & Kipp, H. (2008). Academic achievement over 8 years among children who met modified criteria for Attention-Deficit Hyperactivity Disorder at 4-6 years of age. *Journal of Abnormal Child Psychology 36 (3),* 399-410.

Maxwell, J.A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review, 62*, 279-299.

Maxwell, J.A. (2004). Causal explanation, qualitative research and scientific inquiry in education. *Educational Researcher 33 (1),* 3-11.

Maxwell, S.E. & Delaney, H.D. (2004). *Designing experiments and analysing data: A model comparison perspective. (2nd edition).* Mahwah, NJ: Erlbaum.

May, K. & Hittner, J.B. (2010). Reliability and validity of gain scores considered graphically. *Perceptual and Motor Skills 111 (2),* 399-406.

Merry, S.N., McDowell, H.H., Hetrick, S.E., Bir, J.J. & Muller, N. (2004) Psychological and/or educational interventions for the prevention of depression in children and adolescents. *The Cochrane Database of Systematic Reviews 2004, Issue 2.* Wiley and Sons Ltd.

Mertens, D. (2010). *Research and Evaluation in Education and Psychology: Integrating diversity with quantitative, qualitative and mixed methods. (3rd edition).* Thousand Oaks, CA: Sage.

Miller, A. & Leyden, G. (1999). A coherent framework for the application of psychology in schools. *British Educational Research Journal, 25 (3),* 389-400.

Miller, A. & Todd, Z. (2002). Educational psychology and difficult behaviour in schools: conceptual and methodological challenges for an evidence-based profession. *Educational and Child Psychology, 19 (3),* 82-95.

Miller, L.D., Laye-Gindhu, A., Liu, Y., March, J.S., Thordarson, D.S. & Garland, E.J. (2011). Evaluation of a preventive intervention for child anxiety in two randomized attention-control school trials. *Behaviour Research and Therapy 49 (5)*, 315-323.

Moore, D & McCabe, D. (1993*). Introduction to the Practice of Statistics (2nd edition).* New York: Freeman.

Morrison, K.R.B. (2002). *School Leadership and Complexity Theory.* London: RoutledgeFalmer.

Morton, J. & Frith, U. (1995). Causal modeling: a structural approach to developmental psychopathology in D. Cichetti and D.J. Cohen (eds). *Manual of Developmental Psychopathology*. NY: Wiley.

Mostert, J & Loxton, H. (2008). Exploring the Effectiveness of the FRIENDS program in reducing Anxiety Symptoms Among South African Children. *Behaviour Change 25(2),* 85-96.

Multon, K.D., Brown, S.D. & Lent, R.W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38 (1),* 30-38.

Muris, P. (2002). Relationships between self-efficacy and symptoms of anxiety disorders and depression in a normal adolescent sample. *Personality and Individual Differences 32 (2),* 337-348.

Neil, A.L. & Christensen, H. (2009). Efficacy and effectiveness of school-based prevention and early intervention programs for high anxiety. *Clinical Psychology Review, 29,* 208-215.

Nevo, G.A. & Manassis, K. (2009). Outcomes for treated anxious children: a critical review of long term follow-up studies. *Depression and Anxiety, 26 (7),* 650-660.

Newman, T. (2004). *What works in building resilience?* Essex: Barnado's.

Nijmeijer, J.S., Minderaa, R.B., Buitelaar, J.K., Mulligan, A., Hartman, C.A. & Hoekstra, P.J. (2008). Attention-deficit hyperactivity disorder and social dysfunctioning. *Clinical Psychology Review 28 (4)*, 692-708.

Norman, G. (2010). Likert-scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15,* 625-632.

O'Connor, S., Carney, T., House, E., Ferguson, E. & O'Connor, R. (2010). *PI-ED: The Paediatric Index of Emotional Distress Manual.* London: GL Assessment Ltd.

O'Connor, T. & Creswell, C. (2005). Cognitive behaviour therapy in developmental perspective. In P. Graham, (Ed), *Cognitive Behaviour Therapy for Children and Families (Second Edition)*. Cambridge University Press.

Oakes, J.M. & Feldman, H.A. (2001). Statistical Power for nonequivalent pretest-posttest designs: The impact of change score versus ANCOVA models. *Evaluation Review 25 (1),* 3-28.

Office for Standards in Education (Ofsted, 2010). *Evaluation schedule of* 154*judgements for schools inspected under section five of the Education Act 2005, from September 2009.* Retrieved 18.05.11 from http://www.ofsted.gov.uk/Ofsted-home/Forms-and-guidance/Browse-all-by/Other/General/Framework-for-the-inspection-of-maintained-schools-in-England- from-September-2009.

Ollendick, T. H., Shortt, A. L., & Sander, J. B. (2005). Internalizing disorders in childhood and adolescence. In J. Maddux & B. Winstead (Eds.), *Psychopathology: Foundations for a contemporary understanding* (pp. 353–376.). NJ: Erlbaum Associates.

Oosterwegel, A., & Oppenheimer, L. (1993). *The self-system: Developmental changes between and within self-concepts.* Hillsdale, NJ: Erlbaum.

Outhwaite, W. (1998). Realism and Social Science. In M.S. Archer, R. Bhaskar, A. Collier, T. Lawson & A. Norrie (eds). *Critical Realism: essential readings.* London: Routledge.

Oxford English Dictionary (2002). J. Pearsall & B. Trumble (eds). Oxford: Oxford University Press.

Pajares, F. (1996a). *Assessing self-efficacy beliefs and academic outcomes: The case for specificity and correspondence.* Paper presented at the Annual Meeting of the American Educational Research Association, New York.

Pajares, F. (1996b). Self-efficacy beliefs in academic settings. *Review of Educational Research 66 (4)*, 543-578.

Pajares, F. & Schunk, D.H. (2001). Self-beliefs and school success: self-efficacy, self-concept and school achievement. In R.G Riding and S.G. Rayner (eds). *International perspectives on individual differences (Volume 2: Self-Perception).* Connecticut: Ablex.

Pawson, R. & Tilley, N. (1997). *Realistic Evaluation.* London: Sage.

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies and typologies: Horses for Courses. *Journal of Epidemiology and Community Health, 57*, 527-529.

Petticrew, M. & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide.* Oxford: Blackwell.

Piers, E.V. & Harris, D.B. (1969). *The Piers-Harris children's self-concept scale.* Nashville, TN: Counselor Recordings and Tests.

Pine, D., Cohen, P., Gurley, D., Brook, J. & Ma, Y. (1998). The risk of early adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders. *Archives of General Psychiatry, 55*, 56-64.

Prins, P.J.M. (2001). Affective and cognitive processes and the development and maintenance of anxiety and its disorders. In W.K. Silverman & P.D.A. Treffers (Eds). *Anxiety disorders in children and adolescents.* Cambridge: Cambridge University Press.

Rait, S., Monsen, J.J & Squires, G. (2010). Cognitive Behaviour Therapies and their implications for applied educational psychology practice. *Educational Psychology in Practice 26 (2),* 105-122.

Ramchandani, P., Joughin, C. & Zwi, M. (2001). Evidence-based child and adolescent mental health services: Oxymoron or brave new dawn? *Child Psychology and Psychiatry Review 6 (2)*, 59-64.

Rayner, S.G. (2001). Aspects of the Self as Learner: Perception, concept and esteem in R.J Riding and S.G. Rayner (eds). *International Perspectives*

on Individual Differences: Volume 2, Self-Perception. Westport, Connecticut: Ablex Publishing.

Reichardt, C.S. & Rallis, S.F. (1994). *The qualitative-quantitative debate: New perspectives.* San Francisco: Jossey-Bass.

Reichardt, C.S., & Mark, M.M. (2001). Nonequivalent group designs. In N.J. Smelser and P.B. Baltes. *International Encyclopedia of the Social and Behavioral Sciences*. Elsevier Science Ltd.

Robson, C. (2002). *Real World Research (2nd edition).* Oxford: Blackwell.

Rogosa, D. (1988). Myths about longitudinal research. In. K. W. Schaie, R.T. Campbell, W.M Meredith & S.C. Rawlings (eds). *Methodological issues in aging research, (pp171-209).* New York, NY: Springer.

Rogosa, D.R. & Willett, J.B. (1983). Demonstrating the Reliability of the Difference Score in the Measurement of Change. *Journal of Educational Measurement 20 (4),* 335-343.

Rose, H., Miller, L & Martinez, Y. (2009). 'FRIENDS For Life: The Results of a Resilience-Building, Anxiety prevention Program in a Canadian Elementary School. *Professional School Counselling 12(6)*, 400-407.

Sapp, M. (2004). Cognitive-behavioral theories of counseling: Traditional and non-traditional approaches. Springfield, IL: Charles C. Thomas.

Sayer, A. (1992). *Method in Social Science: A realist approach. (2nd edition).* London: Routledge.

Schoenfeld, N.A., & Janney, D.M. (2008). Identification and Treatment of Anxiety in Students with Emotional or Behavioural Disorders: A review of the Literature. *Education and Treatment of Children, 31(4),* 583-610.

Schoon, I. (2006). Selection, causation and cumulative risk effects. in I. Schoon, *Risk and resilience: adaptations in changing times.* UK: Cambridge University Press.

Schunk, D. H. & Pajares, F. (2009). Self-Efficacy Theory in K.R. Wentzel and A. Wigfield (eds). *Handbook of motivation in school.* New York: Routledge.

Scott, S. (2002). Classification of psychiatric disorders in childhood and adolescence: building castles in the sand? *Advances in Psychiatric Treatment 8,* 205-213.

Seiffge-Krenke, I. (2000). Causal links between stressful events, coping style and adolescent symptomatology. *Journal of Adolescence 23 (6)*, 675-691.

Shadish, W.R., Cook, T.D & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalised causal inference*. Boston: Houghton: Mifflin.

Shaffer, D.R. (1996). *Developmental psychology: Childhood and adolescence (4th edition)*. California: Brooks Cole.

Shavelson, R. & Bolus, R. (1982). Self-concept: The inter-play of theory and methods. *Journal of Educational Psychology 74*, 3-17.

Shavelson, R., Hubner, J. & Stanton, G. (1976). Self-concept: validation of construct interpretations. *Review of Educational Research, 46,* 407-441.

Shunsen, C. & Guiqing, Y. (2009). The relationship of test anxiety and academic achievement: The role of the moderated mediator. Psychological Research, abstract downloaded from http://en.cnki.com.cn on 01.04.11.

Slavin, R.E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researchers 31 (7)*, 15-22.

Smiley, P. & Dweck, C.S. (1994). Individual differences in achievement goals among young children. *Child Development 65*, 1723-1743.

Smolkowski, K. (2010) *Gain Score Analysis*. Retrieved from http://www.ori.org/~keiths/Files/Tips/Stats_GainScores.html on 12.04.11

Spence, S.H. (1997). Structure of anxiety symptoms among children: a confirmatory factor-analytic study. *Journal of Abnormal Child Psychology 106,* 280-297.

Spence, S.H. & Shortt, A.L. (2007). Research review: Can we justify the widespread dissemination of universal school interventions for the prevention of depression among children and adolescents? *Journal of Child Psychology and Psychiatry, 48 (6),* 526-542.

Squires, G. (2001). Using cognitive behavioural psychology with groups of pupils to improve self-control of behaviour. *Educational Psychology in Practice 17 (4),* 317-335.

Squires, G. (2010). Countering the argument that educational psychologists need specific training to use cognitive behavioural therapy. *Emotional and Behavioural Difficulties 15 (4)*, 279-294.

Stallard, P. (2002). Cognitive Behaviour Therapy with children and young people: A selective review of key issues. *Behavioral and Cognitive Psychotherapy 30,* 297-309.

Stallard, P. (2009). Cognitive behaviour therapy with children and young people. In H. Beinart, P. Kennedy & S. Llewelyn, (Eds). *Clinical Psychology in Practice*. UK: Blackwell.

Stallard, P. (2010). Mental health prevention in UK classrooms: the FRIENDS anxiety mental health programme. *Emotional and Behavioural Difficulties, 15 (1),* 23-35.

Stallard, P., Simpson, N., Anderson, S., Carter, T., Osborn, C & Bush, S. (2005). An Evaluation of the FRIENDS programme: A Cognitive Behaviour Therapy Intervention to promote Emotional Resilience. *Archives of Disease in Childhood 90*, 1016-1019.

Stallard, P., Simpson, N., Anderson, S. & Goddard, M. (2008). The FRIENDS emotional health prevention programme 12 month follow-up of a universal UK school-based trial. *European Child and Adolescent Psychiatry 17(5)*, 283-289.

Stallard, P., Simpson, N., Anderson, S., Hibbert, S & Osborn, C. (2007). The FRIENDS emotional health programme: Initial findings from a school-based project. *Child and Adolescent Mental Health 12(1)*, 32-37.

Stallard, P., Simpson, N. & Green, S. (In press). The FRIENDS preventive cognitive behaviour therapy programme: Experience of implementing FRIENDS in school sin the UK. In *Prevention in the school setting*. A. Matsopoulos (ed). Atrapos Publishing House.

Stevens, J.P. (1999). *Intermediate Statistics: A Modern Approach (2nd edition)*. Mahwah, NJ: Erlbaum.

Tew, J. (2005). *Social perspectives in mental health: developing social models to understand and work with mental distress*. London: Jessica Kingsley.

Trochim, W.M.K. (2006). *Nonequivalent Groups Analysis*. Downloaded from http://www.socialresearchmethods.net/kb/statnegd.php on 19.03.11.

University of Nottingham (2009). *Code of Research Conduct and Research Ethics. (Version 2.0)*. University of Nottingham Research Committee.

Usher, E.L. & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78 (4),* 751-796.

Van Ameringen, M., Mancini, C., & Farvolden, P. (2003). The impact of anxiety disorders on educational achievement. *Anxiety Disorders, 17*, 561–571.

Velleman, P.F. & Wilkinson, L. (1993). Nominal, ordinal, interval and ratio typologies are misleading. *American Statistician 47*, 65-72.

Watson, W. (2005). Rethinking the mood and anxiety disorders: a quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology, 114*, 522-536.

Weare, K. (2004). *Developing an emotionally literate school*. London: Sage.

WHO (1992). International Classification of Diseases, Tenth revision. Geneva, Switzerland: WHO.

WHO (2001). *Strengthening mental health promotion. (Factsheet no. 220).* Geneva: WHO

WHO (2004). Promoting mental health: Concepts, emerging evidence, practice (Summary report). Geneva: WHO.

Wigelsworth, M., Humphrey, N., Kalambouka, A. & Lendrum, A. (2010). A review of key issues in the measurement of children's social and emotional skills. *Educational Psychology in Practice 26 (2),* 173-186.

Williams, S.L. (2005). Anxiety Disorders in J.E. Maddux & B.A. Winstead, Psychopathology: *Foundations for a Contemporary Understanding*. NJ: Lawrence Erlbaum.

Willig, C. (2001). *Introducing qualitative research in psychology: Adventures in theory and method.* Buckingham: Open University Press.

Wolpert, M., Fuggle, P., Cottrell, D., Fonagy, P., Phillips, J., Pilling, S., Stein, S. & Target, M. (2006). *Drawing in the evidence: advice for mental health professionals working with children and adolescents (Second Edition).* London: CAMHS Evidenced-Based Practice Unit.

Wood, J. (2006). Effect of anxiety reduction on children's school performance and social adjustment. *Developmental Psychology 42(2),* 345-349.

Wright, D.B. (2006). Comparing groups in a before-after design: When t test and ANCOVA produce different results. *British Journal of Educational Psychology, 76*, 663-675.

Zahn-Waxler, C., Klimes-Dougan, B. & Slattery, M.J. (2000). Internalizing problems of childhood and adolescence: prospects, pitfalls and progress in understanding the development of anxiety and depression. *Development and Psychopathology, 12,* 443-466.

Zimmer-Gembeck, M.J. & Skinner, E.A. (2011). The development of coping across childhood and adolescence: An integrative review and critique of research. *International Journal of Behavioral Development, 35 (1)*, 1-17.

Zimmerman, B.J., Bandura, A. & Martinez-Pons (1992). Self-motivation for academic attainment: the role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*, 663-676.

Secondary References

Ahern, K.J. (1999). Ten tips for reflexive bracketing. Qualitative Health Research 9, 407-411. Cited in C. Robson (2002). *Real World Research (2nd edition)*. Oxford: Blackwell.

Beck, A. (1979). Cognitive Therapy and the Emotional Disorders. New York: Penguin Books. Cited in S. Rait, J. J. Monsen, & G. Squires, (2010). Cognitive Behaviour Therapies and their implications for applied educational psychology practice. *Educational Psychology in Practice 26 (2),* 105-122.

Coe, (2000). What is an effect size? Durham: CEM Centre, University of Durham: www.cemcentre.org/ebeuk/research/effectsize/ESbrief.htm. Retrieved 07.01.05. Cited in L. Cohen, L. Manion and K. Morrison. *Research Methods in Education 6th edition.* London: Routledge.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum. Cited in L. Cohen, L. Manion & K. Morrison. (2007). *Research Methods in Education (6th edition)*. London: Routledge.

Ellis, A. (1957). Rational psychotherapy and individual psychology. *Journal of Individual Psychology*, 13, 38-44. Cited in S. Rait, J. J. Monsen, & G. Squires, (2010). Cognitive Behaviour Therapies and their implications for applied educational psychology practice. *Educational Psychology in Practice 26 (2),* 105-122.

James, W. (1890*). The principles of psychology*. New York: Henry Holt. Cited in S.G. Rayner. Aspects of the Self As a Learner: perception, concept and esteem. *International Perspectives on Individual Differences: Volume 2, Self-Perception*. Westport, Connecticut: Ablex Publishing.

Lang, (1979). A bio-informational theory of emotional imagery. Psychophysiology 6, 495-511. Cited in T.H. Ollendick, A.L. Shortt, & J.B. Sander (2005). *Internalizing disorders in childhood and adolescence*.

Lord, F.M. (1967). A paradox in the interpretation of group comparisons. Psychological Bulletin, 72, 304-305. Cited in D.B. Wright. Comparing groups in a before-after design: when *t* test and ANCOVA produce different results. *British Journal of Educational Psychology 76*, 663-675.

Nunnally, J.C. (1978). Psychometric Theory, (2nd edition). McGraw-Hill, New York. Cited in S. O'Connor, T. Carney, E. House, E. Ferguson and R. O'Connor., (2010). *PI-ED: The Paediatric Index of Emotional Distress Manual.* London: GL Assessment Ltd.

Stevens, S.S. (1946). On the theory of scales of measurement. Science, 103, 677-680. Cited in P.F. Velleman & L. Wilkinson, (1993). Nominal, ordinal, interval and ratio typologies are misleading. *American Statistician 47*, 65-72.

References for Additional Measures Described in this Study

Achenbach, T.M. (1991). *Manual for the child behaviour checklist/ 4-18 and 1991 profile.* Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth and Families.

Battle, J. (1992). *Culture-free self-esteem inventories, 2nd edition.* Austin, Texas: Pro-Ed.

Beck, J.S., Beck, A.T., Jolly, J.B. & Steer, R.A. (2005). *Beck Youth Inventories: Second Edition for Children and Adolescents* (BY-II). PsychCorp, San Antonio, TX.

Brodzinsky, D.M., Elias,M.J., Steiger, C., Simon, J., Gill, M. & Clark Hitt, J. (1992). The Coping Scale for Children and Youth. *Journal of Applied Developmental Psychology* 13, 195-214.

Connell, J.P. (1985). A new multidimensional measure of children's perceptions of control. *Child Development*, 56, 1018-1041.

Harter, S. (1985). Manual for the self-perception profile for children. Denver, Co: University of Denver. Cited in Burden, R. (1998b). Myself As Learner Scale (MALS

Hernández- Gúzman, L. (2003). *Escala de afrontamiento (versión infantil). Proyecto de investigación DGAPA IN-302600, evaluación y categorización de los trastornos de ansiedad en niños y adolescentes.* México, DF.: Universidad Autónoma de México (UNAM), Facultad de Psicología.

Kovacs, M. (1981). *Rating scales to assess depression in school-aged children*. Acta Paedopsychiatry, 46, 305-315.

March, J.S. (1997). *The Multidimensional Anxiety Scale for Children.* North Tonawanda, NY: Multi-health systems.

Piers, E. V. (1984). *Piers-Harris children's self-concept scale: Revised manual.* Los Angeles: Western Psychological Services.

Reynolds, C.R. & Richmond, B.O. (1978). What I think and feel: A revised measure of children's manifest anxiety. *Journal of Abnormal Child Psychology* 6, 271-280.

Schwab-Stone, M.E., Shaffer, D., Dulcan, M.K., Jensen, P.S., Fisher, P., Bird, H.R., Goodman, S.H., Lahey, B.B., Lichtman, J.H., Canino, G., Rubio-Stipec, M & Rae, D.S. (1996). Criterion Validity of the NIMH Diagnostic Interview

Schedule for Children Version 2. (DISC-2.3). Journal of the American Academy of Child and Adolescent Psychiatry, 35, 878-888.

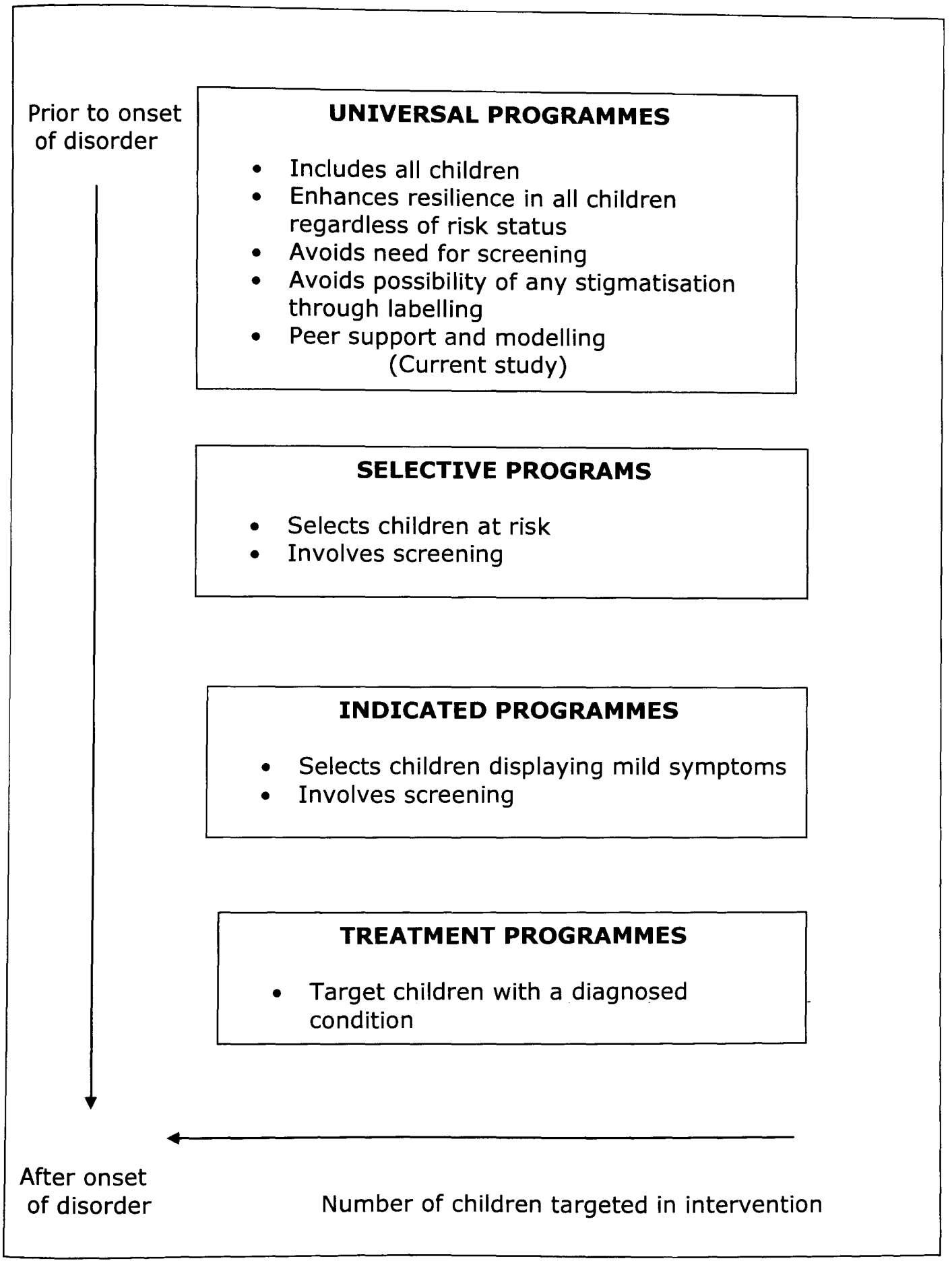
Silverman, W.K. & Albano, A.M. (1997). *The anxiety disorders interview schedule for DSM-IV*. San Antonio: The Psychological Corporation.

Spence, S.H. (1997). *The Spence Children's Anxiety Scale (SCAS).* www.scaswebsite.com

Zigmond, A.S.. and Snaith, R.P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67, 361-70.

8 APPENDICES

Appendix 1a: The Intervention Continuum (Adapted from Lowry-Webster, Barrett and Dadds, 2001).



Appendix 2a: Search Strategy for Literature Review (Sections 2.1-2.4)

A non-systematic electronic search was conducted using the online databases Google scholar, EBSCO, PSYCINFO, MEDLINE and Cochrane Database of Systematic Reviews. The University of Nottingham online catalogue was used to follow up references, including online journals and book chapters. Texts recommended through research supervision sessions were also consulted. The researcher aimed to be as comprehensive as possible, using 'Whole Text' searches and refining the combination of key words within each area of investigation until no new items were found. The search was initially broad, to include important historical sources, but was then restricted to 2000 onwards to identify the most relevant, contemporary literature relating to the identified theories.

Key words:

1. mental health
2. risk factors
3. protective factors
4. coping
5. anxiety
6. depression
7. emotional distress
8. child*
9. primary school
10. elementary school
11. academic
12. self-concept
13. self-efficacy
14. self-perceptions
15. school

16. achievement
17. performance
18. cognitive behaviour*
19. cognitive behavior*
20. programme
21. treatment
22. therapy

Appendix 2b: Systematic search strategy FRIENDS (Section 2.5)

The following electronic databases were searched: Google, Google scholar, FRIENDS website, EBSCO, PSYCINFO, MEDLINE, DARE, Current Controlled Trials and Cochrane. The University of Nottingham Library online catalogue and interlibrary loan service were used to obtain items unavailable from other online sources.

Key words used in this search were:

1. "FRIENDS"
2. "FRIENDS For Life"
3. Program
4. Programme
5. Universal
6. School-based
7. Trial
8. Intervention
9. Barrett

Criteria for included papers:

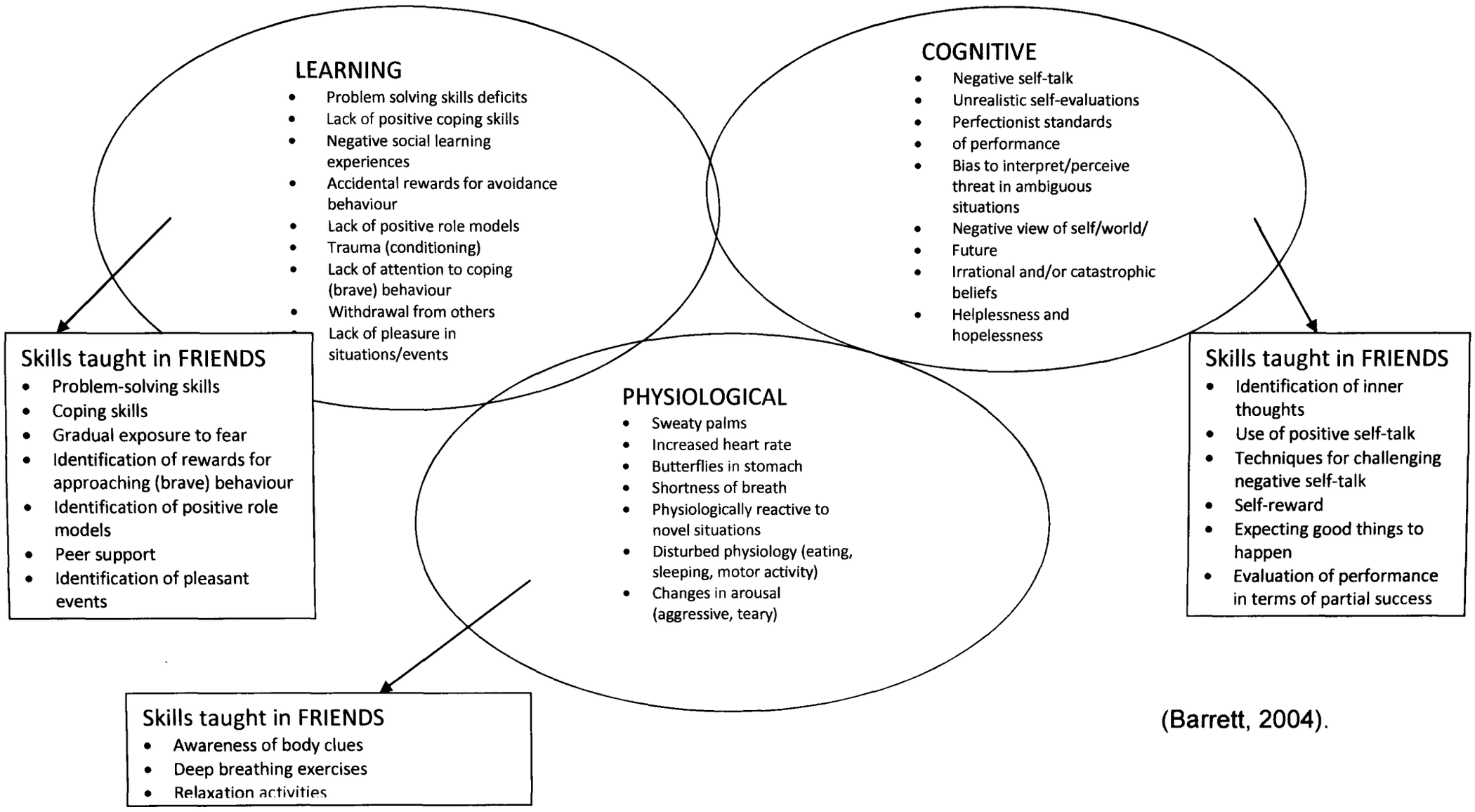
- 'FRIENDS For Life' had to be included as at least one of the evaluated interventions as opposed to exclusively 'Fun Friends' (younger population) or 'FRIENDS For Youth' (adolescent population).
- Designs employing randomised controlled trials or quasi-experiments with a wait-list control were prioritised. Due to the lack of evidence from UK-based research, however, pre/post test only designs with no control/comparison group were also included.
- The programme was delivered as a universal, school-based intervention, involving whole classes from designated schools.

- The sample derived from a mixed, school-based population with ages ranging between 7 and 13.
- Outcomes included a specific, standardised measure of anxiety/depression.
- The paper involved a level of peer review (published in journal/doctoral thesis).

Excluded papers comprised:

- Studies including only qualitative analysis.
- Studies evaluating FRIENDS as a selective, indicated or individual treatment programme.
- Studies focusing on specific vulnerable subgroups within the general population (eg. samples from particular cultural minorities).
- Studies focusing exclusively on younger (below 7) or older (above 13) age groups.

Appendix 2c: How FRIENDS utilises the tripartite model of anxiety



(Barrett, 2004).

Appendix 2d: Summary tables for universal trials of FRIENDS For Life

Authors Date	Title	Journal	Population	Intervention	Design	Method Quality	Measures	Outcomes
Barrett, P. M. and Turner, C. (2001)	Prevention of anxiety symptoms in primary children: preliminary results from a universal school-based trial	British Journal of Clinical Psychology, 2001, 40, pp399-410.	489 children (aged 10-12 years) from 10 schools in Brisbane, Australia.	Psychologist-led FRIENDS; teacher-led FRIENDS or control group.	Randomised controlled trial	Unit of randomisation = school. Regular integrity checks (88-92%) Groups matched at pre-test. Support teachers and psychs as effective leaders but stat. significance lacks power due to small sample size. Relies on self-report (no parent measures).	Spence Children's Anxiety Scale (1994); Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1978); Children's Depression Inventory (Kovacs, 1981) – all self-report only.	Psych and teacher-led IGs showed sig reductions in anxiety on SCAS. ($F(2,451) = 3.25$; $p < .05$) CDI showed sig increase in teacher-led intervention group (IG) but not psych-led or control group (CG). Psych and teacher-led IGs showed sig reductions in RCMAS. ($F(2,457) = 4.24$; $p < .05$). High anxiety chn in IGs more likely to move from 'at risk' into 'healthy' range but stat sig not detected due to small numbers.
Lowry-Webster, H., Barrett, P.M and Dadds, M.R. (2001)	A Universal Prevention Trial of Anxiety and Depressive Symptomatology in Childhood: Preliminary Data from an Australian study	Behaviour Change, 2001, 18 (1), pp36-50	594 (aged 10-13 years, 314 girls, 280 boys) from 7 schools in Brisbane, Australia.	FRIENDS or comparison group + parent sessions	Randomised controlled trial.	Schools matched for size & demographic info; randomly allocated on class by class basis. Groups matched on SCAS but not on RCMAS or CDI at pre-test. (CG higher). Treatment integrity checked	Spence Children's Anxiety Scale (1994); Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1978); Children's Depression Inventory (Kovacs, 1981)	Chn in IG reported fewer anxiety symptoms at post-test regardless of risk status. SCAS scores sig decreased more for IG than CG. $t(545) = 6.59$, $p<.05$ Greater % remained at risk in CG. (75.3% of IG at risk in pre-test showed sig benefits compared to 42.2%in CG). No sig change on RCMAS for either group. Sig reduction in CDI scores

Authors Date	Title	Journal	Population	Intervention	Design	Method Quality	Measures	Outcomes
						regularly.		for high anxiety IG only.
Lowry-Webster, H.M., Barrett, P.M & Lock, S. (2003)	A Universal Prevention Trial of Anxiety Symptomatology during Childhood: results at 1 year follow-up.	Behaviour Change, 2003, 20 (1) pp25-43	As above	As above – 1 year follow-up	Follow up to Lowry-Webster, H., Barrett, P.M and Dadds, M.R. (2001)	Low response rate to parent CBCL. Use of diagnostic interview a strength although only conducted at 12 month FU on chn with high anxiety and depression scores. Teacher reports omitted.	As above +: Anxiety Disorders Interview Schedule for Children (ADIS-C; Silverman and Albano 1997); Child Behaviour Checklist, CBCL-revised; Achenbach and Edelbrock, 1991). [Treatment Acceptability measures were also administered].	SCAS scores for IG lower than CG at 12 months. High anxiety group maintained their lower scores. Sig relationships between risk status and treatment group found, favouring IG. 85% of high anxiety group were diagnosis free at 12 months compared to 31.2% in CG. Conclusion: Intervention gains largely maintained over 12 months according to self-report & diagnostic interview.
Lock, S. & Barrett, P.M. (2003)	A Longitudinal Study of Developmental Differences in Universal Preventive Intervention for Child Anxiety	Behaviour Change 2003, 20 (4), pp183-199.	733 children in Grade 6 (aged 9-10, n=336) and grade 9 (aged 14-16, n=401) from diverse socio-economic backgrounds in Brisbane.	FRIENDS for Life IG and CG.	Randomised Controlled Trial (school as unit of randomisation. Participants stratified into 'at risk' and 'healthy' groups.	Integrity checks completed but results not reported. Pre-intervention group differences not controlled for according to Barrett et al. (2006).	Spence Children's Anxiety Scale (1994); Children's Depression Inventory (Kovacs, 1981); Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1978); Coping Scale for Children and	Sig reductions in anxiety and depression reported by IG and CG at post-test. ($F(6, 23) = 45.49$, $p<0.001$). IG showed greater anxiety reductions at post test and 12 month FU than CG. Grade 6 reported sig higher levels of anxiety pre-intervention and at post-test but greater reductions than Grade 9 at 12 month follow-up. No differences between IG and CG in 'at risk' group changes but this may be due

Authors Date	Title	Journal	Population	Intervention	Design	Method Quality	Measures	Outcomes
							Youth (Brodzinsky et al., 1992); Anxiety disorder Interview schedule for Children – IV. (ADIS-C-IV; Silverman and Albano, 1996). Self-reports only.	to attrition patterns. Increased cognitive-behavioural problem-solving in Grade 9 and reduced cognitive-behavioural avoidance in Grade 6 but these effects disappeared by 12 month FU.
Barrett, P.M., Farrell, L.J., Ollendick, T.H & Dadds, M. (2006).	Long-Term Outcomes of an Australian Universal Prevention trial of Anxiety and Depression Symptoms in Children and Youth: An Evaluation of the FRIENDS Program.	Journal of Clinical Child and Adolescent Psychology, 2006, 35 (3), pp403-411	669 of original sample reported in Lock, S. & Barrett, P.M. (2003).	As reported in Lock, S. & Barrett, P.M. (2003)	12, 24 and 36 month follow-up to Lock, S. & Barrett, P.M. (2003)	No clustering effect of schools.	Spence Children's Anxiety Scale (1994); Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 1978); Children's Depression Inventory (Kovacs, 1981)	IG reductions in anxiety maintained for students in Grade 6, with significantly greater reductions at LT FU. No sig group differences in Grade 9 – authors interpret that this supports Grade 6 as optimal time for early intervention. A significant Time x IG x gender effect was found with girls in IG reporting sig lower anxiety at 12 and 24 month FU but not at 36 months in comparison to CG – authors conclude intervening in primary years produces positive short term outcomes for girls at universal level. Prevention effect demonstrated with sig fewer high risk students at 36 month FU in the IG compared to CG.
Barrett, P.M., Lock, S. & Farrell, L.J. (2005).	Develop-mental Differences In Universal	Clinical Child Psychology and Psychiatry,	692, Grade 6 aged between 8-9 (n=293) and	FRIENDS For Life IG and CG.	Randomised controlled trial (school as unit of	Integrity checks revealed 88.8-95.6% concordance	Spence Children's Anxiety Scale (1994);	Post-test results indicated sig reductions in anxiety ($F(2, 1.93) = 7.10$; $p < 0.001$) and depression ($F(2,1.97) = 5.37$;

Authors Date	Title	Journal	Population	Intervention	Design	Method Quality	Measures	Outcomes
	Preventive Interven-tion for child anxiety	2005, 10 (4), pp539-555	Grade 9, aged 14-16 (n=399) from Brisbane, Australia, diverse socio-economic status. IG n=423 CG n=269		randomisation) with 12 month follow-up. Sample stratified into low, moderate and high risk groups for anxiety.	between session and manual content. Poor attendance at parent workshops.	Children's Depression Inventory (Kovacs, 1981) – self-report only.	p<.05) across high and moderate risk groups in IG and CG, but no sig. difference between groups. Sustained at 12 month FU but sig greater reductions in high and moderate risk groups in IG, (F(1,543)=7.29; p<.05). At post-test, Grade 6 scores showed significant reductions compared with Grade 9, (F(2, 1.93) = 13.066; p< .001) although both groups showed equal reductions at 12 month FU.
Stallard, P., Simpson, N., Anderson, S., Carter, T., Osborn, C. & Bush, S. (2005)	An Evaluation of the FRIENDS programme: a cognitive behaviour therapy intervention to promote emotional resilience.	Archives of Disease in Childhood, 2005, 90, pp1016-1019.	213 children aged 9-10 years from 6 primary schools in southwest England.	FRIENDS 10 sessions delivered by school nurses.	Uncontrolled pre and post-test study.	No control group Clinical significance not assessed. No long term follow-up.	Spence Children's Anxiety Scale (1994); Culture Free Self-esteem Questionnaire Form B, (Battle, 1992). Qualitative questionnaire developed by participation worker from Children's Society. Self-report only.	Post-test data revealed sig lower rates of anxiety (t=2.950, p=0.003) and sig improved levels of self-esteem. (t=2.950, p=0.002). Post-test assessments for high risk group revealed sig increase in self-esteem (t=4.789, p= 0.0001) and sig decrease in anxiety (t= 2.362, p= 0.023). Status of 60% of children in high risk group positively changed. Qualitative analysis of acceptability revealed 81% thought programme was fun; 77.4% would recommend to a friend; 72.8% thought they had learned new skills and 41.1% had helped someone else with their new skills.
Stallard, P., Simpson, N.,	The FRIENDS	Child and Adolescent	106 children (60 boys, 46	FRIENDS 10 sessions	Uncontrolled pre and post-	Monthly supervision group	Spence Children's	No sig change from T1-T2 – ie. indicating stable anxiety

Authors Date	Title	Journal	Population	Intervention	Design	Method Quality	Measures	Outcomes
Anderson, S., Hibbert, S. & Osborn, C. (2007)	Emotional Health programme: Initial findings from a School-based Project.	Mental Health, 2006, 12 (1); pp32-37	girls) aged 9-10 from 3 schools in Bath and NE Somerset. (1 school with high rate of EBD; 1 from severely deprived area and 1 rural).	delivered by school nurses.	test study. Measures taken T1 6 months before; T2 upon starting and T3 3 months after finishing programme.	but no details of treatment integrity. Small sample Single cohort design.	Anxiety Scale (1994); Culture Free Self-esteem Questionnaire Form B, (Battle, 1992).	and self-esteem prior to intervention. Significant change for total anxiety (F= 5.84, p=0,003) and self-esteem (F=2.98, p=0.052) across time. Reduction in anxiety for 'high anxiety' group was sig. (F=5.30, p=0.011) and increase in self-esteem for low self esteem group was significant (F=5.78, p=0.043).
Stallard, P., Simpson, N., Anderson, S & Goddard, M. (2008).	The FRIENDS emotional health prevention programme: 12 month follow-up of a universal UK school-based trial.	European Child and Adolescent Psychiatry 2008, 17 (5), pp283-289.	As reported by Stallard, P., Simpson, N., Anderson, S., Hibbert, S. & Osborn, C. (2007).	As reported by Stallard, P., Simpson, N., Anderson, S., Hibbert, S. & Osborn, C. (2007).	As reported by Stallard, P., Simpson, N., Anderson, S., Hibbert, S. & Osborn, C. (2007) – 12 month follow-up.	Small sample size No comparison group	Spence Children's Anxiety Scale (1994); Culture Free Self-esteem Questionnaire Form B, (Battle, 1992).	Sig effect over time for total self-esteem (F(3,323)=6.55, p=0.0001) and anxiety (F(3,323) = 8.58, p=0.0001). No sig differences between T3and T4 analyses – ie. Benefits maintained at 12 months. Of 9 children identified as high risk at T2, 6 moved into low risk by 12 months. No low risk moved into high risk – preventive effect.
Mostert, J. & Loxton, H. (2008)	Exploring the Effectiveness of the FRIENDS program in Reducing Anxiety Symptoms Among South African Children	Behaviour Change, 2008, 25 (2), pp 85-96.	46 (n=25 in IG; n=21 in CG), 12 year old South African children (30 girls, 36 boys) from low socioeconomic background Ad hoc convenience sample.	FRIENDS	Quasi-experimental non-equivalent control group design with 4 and 6 month follow-up. (CG received intervention after Time 3).	Non-random design. Groups matched for age, gender, anxiety at pre-test. Small sample size.	Spence Children's Anxiety Scale (1994);	IG SCAS scores showed sig decrease between Time 1 and Time 3 and Time 1 and Time 4, (F (3) = 11.46, p=0) but not between Time 1 and Time 2. The decline in scores for the CG was not sig. No sig between groups differences on SCAS at any time point.

Authors Date	Title	Journal	Population	Intervention	Design	Method Quality	Measures	Outcomes
Gallegos, J. (2008).	Preventing childhood anxiety and depression: testing the effective-ness of a school-based program in Mexico.	The University of Texas at Austin, 2008, 196 pages, 3341564. https://www.lib.utexas.edu/etd/d/2008/gallegosd87338/gallegosd87338.pdf	1,030 4th and 5th grade students from 8 schools in a northern city in Mexico (IG n=534; CG n=496)	AMISTAD (Spanish version of FRIENDS) IG and CG.	Quasi-experimental non-equivalent comparison group design. School was unit of random assignment. Participants allocated to one of 4 non-overlapping groups – chn anxiety diagnosis- free and non learning difficulty (LD); chn at risk for anxiety and non LD; chn at risk with LD and chn diagnosis free with LD. 6 month follow-up.	Spillover effects controlled by school being unit of randomisation. Schools matched on socioeconomic levels; individuals stratified according to anxiety risk to improve causal inference. Groups matched at pretest. Treatment integrity revealed moderate to good results. Outcomes rely mainly on self-report measures; few assessments targeted positive, strength based outcomes. Attendance data not available.	Spanish version of Spence Children's Anxiety Scale (1997); Spanish version of Children's Depression Inventory (Kovacs, 1981); Cuestionario de Afrontamiento (Hernandez-Guzman, 2003) – measures coping skills; Spanish version of Pier-Harris Children's Self-Concept scale (CSCS: Piers, 1984); Spanish version of Anxiety Disorder Interview schedule for children (ADIS-C-IV, Silverman and Albano, 1996); Spanish version of Child Behaviour Checklist, Parent Version (CBCL, Achenbach &	Statistically significant improvements of small impact for the overall sample and for children diagnosis-free and non-LD, in that those receiving the programme decreased the severity of their depressive symptoms and increased their proactive coping skills. For children already showing risk for anxiety and/or learning difficulty, the program did not produce meaningful changes. Children at risk of depression decreased by 2.6% in IG and increased by 5.4% in CG. (Preventive effect). No significant increase in self-concept was found for the children with LD, (only group tested on this measure).

Authors Date	Title	Journal	Population	Intervention	Design	Method Quality	Measures	Outcomes
							Rescorla, 2001). Self-report only.	
Rose, H., Miller, L. & Martinez, Y. (2009)	'FRIENDS For Life': The Results of a Resilience-Building, Anxiety Prevention Program in a Canadian Elementary School.	Professional School Counselling, 2009, 12 (6), pp400-407	52 students aged 8-9 years (IG n=26; CG n=26) from an urban elementary school in western Canada.	FRIENDS For Life delivered in 8 sessions.	Non randomized control group pre-test/post-test design with 2 groups.	Small sample size. Control group had sig lower rates of anxiety at pre-test. No parent or teacher evaluations included.	Multi-dimensional Anxiety Scale for Children (MASC; March 1997). Pupil and parent perceptions of programme gained through questionnaires	All children reported lower rates of anxiety at post-test but no significant within group or between group differences found.

Appendix 3a: Letter to School

Please ask for: Julie Paul
Tel: 01604 630082
Our ref: JP/PD
Your ref:
Date: 17th June 2010

Dear Mr ,

An opportunity has arisen to take part in some research evaluating the 'FRIENDS For Life' programme. This is an intervention that aims to reduce children's anxiety and enhance their coping skills in a range of situations. The programme has been well received in other schools in Northamptonshire and across the country, and research suggests that it can have very positive effects on children's well-being. The present study is a doctoral thesis and contributes to both the county and national evaluation of targeted mental health programmes. It also aims to find out whether FRIENDS improves children's views about themselves as learners which might suggest that it could have a beneficial impact on academic attainment.

How would this help your school?

Research has indicated that FRIENDS has a positive effect on children's anxiety, self-esteem and behaviour. It also suggests that the techniques taught in the programme will help the children to develop better coping strategies which enable them to function more successfully at school and at home. As well as taking measures of anxiety, data will be gathered on children's self-perceptions as learners and on teacher perceptions of pupil strengths and difficulties. This will provide you with very useful information to report in your school evaluation.

What will the project look like?

Your learning mentor who has attended the FRIENDS Training will deliver the 10 week programme with my support in one of the Year 5 classes. The parallel class will act as a control group in the study and will also participate in the measures. They will complete the FRIENDS programme later in the academic year.

Northamptonshire County Council
Children and Young People's Service - Northampton Area
Springfield, Cliftonville.
Northampton. NN1 5BE
w. www.northamptonshire.gov.uk
t. 01604 630082
f. 01604 630283

Northamptonshire County Council

Before the Summer holiday, the teachers, learning mentor and I will conduct a planning meeting. Parental consent *for both classes* to participate in the measures will need to be obtained by Monday 19th July; I will provide a letter for your approval and discuss the best method for dissemination. Early in the Autumn term, I will come into school to introduce the project to the pupils and to take the initial set of measures. The programme will then run weekly from September to December when the second set of measures will be taken. Results will be reported back in the Summer Term of 2011.

Allocation of time

The programme is based on ten 1 hour weekly sessions. In addition to this, I would need half an hour to introduce my project to the class, to discuss ethical considerations and then to withdraw children in small groups to complete the scales. I anticipate that this would take no more than one day to complete. It would also be helpful if a session could be provided for the teacher to fill in the Strengths and Difficulties Questionnaires.

Ethical Considerations

Participants will have the right to withdraw from the study at any time and do not need to provide a reason. Data will be securely kept and all final reporting will be anonymous. If any child's scores on the anxiety scale are within the range of concern, this will be brought to the attention of teachers initially, then parents if necessary and further measures taken as required.

If you agree to your school taking part in this study, please sign the consent form attached. If at any time you decide that you do not want to continue to participate you are free to withdraw without qualification.

If you have any further questions please do not hesitate to contact me by phone 01604 630082 or e-mail jpaul@nothamptonahire.gov.uk

Yours sincerely,

Julie Paul, Trainee Educational Psychologist

CONSENT FORM

Investigating the effects of the FRIENDS for Life programme on anxiety and academic self-perceptions.

Researcher: Julie Paul

Supervisor: Anthea Gulliford

School of Psychology, University of Nottingham

Please circle:

Have you read and understood the covering letter?	YES/NO
Have you had the opportunity to ask questions and discuss the study?	YES/NO
Have all arising questions been answered satisfactorily?	YES/NO
Have you received enough information about the study?	YES/NO
Do you understand that your school is free to withdraw from the study:	
at any time?	YES/NO
without having to give a reason?	YES/NO
Do you agree to your school taking part in the study?	YES/NO

"This study has been explained to me to my satisfaction, and I agree to give consent for .....School to take part. I understand that we are free to withdraw at any time."

Signature of Head Teacher: _____

Name: _____

Date: _____

I have explained the study to the Head Teacher and he has agreed for his school to take part.

Signature of researcher: _____
Date: _____

Appendix 3b: Letter to Parents

Please ask for: Julie Paul
Tel: 01604 630082
Our ref: JP/PD
Your ref:
Date:

Dear parent/carer,

As part of the health curriculum next year your child will be taking part in the FRIENDS For Life programme with the rest of their class. This is a ten week programme that teaches young people problem-solving and life-skills. It is currently being used in other schools in Northamptonshire and many have found it to be very helpful. Class____ will be taking part in the autumn term and Class _____ will be taking part in the spring term.

I am a Trainee Educational Psychologist working with Northamptonshire County Council and studying at the University of Nottingham. I am carrying out a study to find out whether FRIENDS is successful at helping the children in local schools. To do this, with your permission, I will be asking the children to answer a few questions about how they feel at the beginning of the programme and again at the end. The results will be reported back during the Summer Term 2011.

All data will be kept confidential unless any of the children's responses give us cause for concern. In this case, you would be invited to come and discuss the matter with myself or the child's class teacher. **You have the right to withdraw you child's data from the study at any point should you wish to and you do not have to give a reason.**

We will be inviting parents to come to _____ classroom on **Wednesday 22nd September at 3.15pm** where you will be able to ask questions and sign the consent form for your child to take part. We will be offering a prize draw for some Tesco vouchers for those who attend.

If you have any further questions or would like to contact me in the future, please see _____

Yours faithfully,

Julie Paul
Trainee Educational Psychologist

Northamptonshire County Council
Children and Young People's Service - Northampton Area
Springfield, Cliftonville.
Northampton. NN1 5BE
w. www.northamptonshire.gov.uk
t. 01604 630082
f. 01604 630283
e. SENNpton@northamptonshire.gov.uk

Northamptonshire County Council

PARENT CONSENT FORM

Investigating the effects of the FRIENDS for Life programme on anxiety and academic self-perceptions.

IF NOT ATTENDING THE PARENT MEETING, PLEASE RETURN THIS FORM TO THE SCHOOL OFFICE BY _____. THANK YOU.

Please circle:

Have you understood the information about the study?	YES/NO
Have you had the chance to ask questions?	YES/NO
Have all of your questions been answered?	YES/NO
Do you understand that you are free to withdraw your child:	
At any time?	YES/NO
Without having to give a reason?	YES/NO
Do you agree to your child taking part in this study?	YES/NO

"This study has been explained to me to my satisfaction and I agree to my child taking part. I understand that I am free to withdraw my child at anytime."

Name of child: _____

Signature of parent/carer: _____

Date: _____

The study has been explained to the above participant's parent/carer and they have agreed to take part.

Signature of researcher: _____

Date: _____

Appendix 3c: Script for Parent's Information and Consent Meeting

1. Learning Mentor provides introduction to FRIENDS programme covering content, purpose, timing etc.

2. My input to cover the following points:

- I am an Educational Psychology student at the University of Nottingham working for Northamptonshire County Council. I've been working with staff and children at _____ since September.

- Carrying out a study to see how well the children respond to FRIENDS in local schools.

- I have asked school whether I can run my study alongside their use of the programme and they have kindly agreed.

- What will this involve? The children answering some questions about how they are feeling now and how they feel at the end of the programme. The sorts of questions we'll be asking are things like 'Do you like having problems to solve?'; 'When you get stuck with your work, can you work out what to do next?' These questions will be very similar to ones that the school ask anyway as part of their evaluations. There are 2 short questionnaires to fill in which should take no more than 10 minutes each.

- Because this study is something extra, we need to have your permission for your child to take part. The purpose of being here today is to answer any questions you might have about the study before agreeing for your child to take part.

- When I write up the report, all of the data will be put together; no names will be used; all of the information will remain confidential; there will be no way of identifying your child individually.

- If any of the children's answers give cause for concern, I would initially raise this with the child's class teacher and then arrange to have a meeting with parents as necessary.

- If you decide at any point while we're carrying out the study that you would prefer your child's questionnaire not to be used, you can withdraw it by letting Mrs _____ know. You will NOT be withdrawing them from the programme itself as the school will be continuing with this anyway. We are just talking now about the measures that I am taking for the study. You do not need to give a reason why you wish to withdraw.

- Does anybody have any questions? If you would prefer to speak to somebody privately, that's fine, please come and find one of us in a moment and we will be happy to answer any queries. Otherwise, please go ahead and fill in the form if you have not done so already. [Read through together and point out the need to circle and sign responses]. We really appreciate the fact that you've given up time for us

so everybody will receive a raffle ticket for some Tesco vouchers to show our appreciation.

Appendix 3d: Timeline for thesis project

p = pilot study m = main study

	Oct 09	Nov 09	Dec 09	Jan 10	Feb 10	Mar 10	Apr 10	May 10	Jun 10	Jul 10	Aug 10	Sep 10	Oct 10	Nov 10	Dec 10	Jan 11	Feb 11	Mar 11	Apr 11	May 11	Jun 11
Pilot study							x	x	x	x											
Sample selection					x (p)																
Formal meeting with senior management					x(p/m)	x(m)															
Initial meeting with parents, teachers, pupils						x (p)						x(m)									
T1 measures							x(p)						x(m)								
FRIENDS intervention							x(p)	x (p)	x(p)	x(p)			x(m)	x(m)	x(m)						
Waiting list control																	x(m)	x(m)	x(m)		
Completion of lit review										x											
T2 measures										x(p)					x(m)						
T3 data for CG																					x
Data analysis											x(p)				x(m)	x(m)	x(m)	x(m)			
Interpretation of data																	x	x			
Complete final write up																		x	x	x	
Hand in																				x	

Appendix 3e: Sample questions from the Paediatric Index of Emotional Distress (O'Connor et al, 2010). Reproduced by kind permission of GL Assessment Ltd.

PI-ED

GL assessment
the measure of potential

Feelings are really important.

Your answers to the questions over the page will help us understand how you feel.

Please read each of the sentences and put a tick (☑) beside the answer that describes you best.

Think about how you have been feeling over the last week when you read each sentence.

There are no right or wrong answers but it is important for you to let us know how you feel.

The sentence below is an example. Please tick the box that best describes how you feel.

I like to play sports.

- ☐ Always
- ☐ A lot of the time
- ☐ Sometimes
- ☐ Not at all

Now turn over the page.

PI-ED

GL assessment
the progression of potential

Name: _____ Date: _____

Age: _____ Please tick: Male ☐ Female ☐

Please tick the box [✓] that best describes how you feel over the last week.

1 I feel sleepy or 'wound up'.
- ☐ Always
- ☐ A lot of the time
- ☐ Sometimes
- ☐ Not at all

2 I feel restless / fidgety as if I have to be on the move.
- ☐ Always
- ☐ A lot of the time
- ☐ Sometimes
- ☐ Not at all

3 I get annoyed easily.
- ☐ Always
- ☐ A lot of the time
- ☐ Sometimes
- ☐ Not at all

4 I can chill-out and feel relaxed.
- ☐ Always
- ☐ A lot of the time
- ☐ Sometimes
- ☐ Not at all

5 I feel good about myself.
- ☐ Always
- ☐ A lot of the time
- ☐ Sometimes
- ☐ Not at all

6 I am lonely.
- ☐ Always
- ☐ A lot of the time
- ☐ Sometimes
- ☐ Not at all

Serial

Appendix 3f: Sample questions from the Myself-As-Learner Scale (Burden, 1998, in Psychology in Education Portfolio, N Frederickson & R.J. Cameron (eds), 1999, NfER Nelson). Reproduced by kind permission of GL Assessment Ltd.

MYSELF AS LEARNER SCALE (MALS)

NFER-NELSON

HOW I SEE MYSELF

Instructions: On the next page you will be given 20 questions to answer. Their purpose is to find out how you see yourself when it comes to learning and school work. Some people see themselves as being very good at learning and doing hard work, but others don't. We want to know what you think about yourself.

This is not a test. There are no right or wrong answers, so please try to answer the questions as truthfully as you can. Your answers will not be shown to anyone else.

First of all we need some information about you.

Name .....

Boy or girl.....

Date of birth .....

Today's date .....

Your age .....

Please read the statements carefully.

If you definitely agree, please put a circle around a

If you agree a bit, but not so strongly, please put a circle around b

If you think that the statement is true about half the time, please put a circle around c

If you don't agree, please put a circle around d

If you strongly disagree, please put a circle around e

	Definitely agree	Agree a bit	True half the time	Don't agree	Strongly disagree
1. I'm good at doing tests.	a	b	c	d	e
2. I like having problems to solve.	a	b	c	d	e
3. When I'm given new work to do, I usually feel confident I can do it.	a	b	c	d	e
4. Thinking carefully about your work helps you to do it better.	a	b	c	d	e
5. I'm good at discussing things.	a	b	c	d	e
6. I need lots of help with my work.	a	b	c	d	e
7. I like having difficult work to do.	a	b	c	d	e
8. I get anxious when I have to do new work.	a	b	c	d	e

Appendix 3g: E-mail from R. Burden regarding standardisation data for the MALS

From: Burden, Robert [mailto:R.L.Burden@exeter.ac.uk]
Sent: 29 January 2010 17:02
To: Julie Paul
Subject: RE: MALS

Dear Julie,
Thank you for your inquiry. Data has been collected on the applicability of the MALS over a wide age range, although this is not yet published. We have found that below the age of nine the scale's reliability can become somewhat suspect. You can try it but you would need to go through every question carefully aloud and ensure that the children understand what is being asked of them and how to respond appropriately. This could be a worthwhile exercise in itself in adding data on the MALS' validity and reliability.
Hope this helps,
Bob Burden.

From: Julie Paul [julie.paul@runbox.com]
Sent: Friday, January 29, 2010 8:45 AM
To: Burden, Robert
Subject: MALS

Dear Professor Burden,

Please could you clarify the age range for which the MALS is currently standardised? In the Psychology in Education Portfolio you indicate that data was being collected over a wider age range and I wondered whether the results of this were available yet. I was hoping to use the scale with a group of 7 year olds but am not sure whether this is too young to provide valid results for formal reporting purposes. I would very much appreciate your feedback.

Kind regards

Julie Paul
Trainee Educational Psychologist

Appendix 3h: Strengths and Difficulties Questionnaire

Strengths and Difficulties Questionnaire T 4-16

For each item, please mark the box for Not True, Somewhat True or Certainly True. It would help us if you answered all items as best you can even if you are not absolutely certain or the item seems daft! Please give your answers on the basis of the child's behaviour over the last six months or this school year.

Child's Name ..... Male/Female

Date of Birth.....

	Not True	Somewhat True	Certainly True
Considerate of other people's feelings	☐	☐	☐
Restless, overactive, cannot stay still for long	☐	☐	☐
Often complains of headaches, stomach-aches or sickness	☐	☐	☐
Shares readily with other children (treats, toys, pencils etc.)	☐	☐	☐
Often has temper tantrums or hot tempers	☐	☐	☐
Rather solitary, tends to play alone	☐	☐	☐
Generally obedient, usually does what adults request	☐	☐	☐
Many worries, often seems worried	☐	☐	☐
Helpful if someone is hurt, upset or feeling ill	☐	☐	☐
Constantly fidgeting or squirming	☐	☐	☐
Has at least one good friend	☐	☐	☐
Often fights with other children or bullies them	☐	☐	☐
Often unhappy, down-hearted or tearful	☐	☐	☐
Generally liked by other children	☐	☐	☐
Easily distracted, concentration wanders	☐	☐	☐
Nervous or clingy in new situations, easily loses confidence	☐	☐	☐
Kind to younger children	☐	☐	☐
Often lies or cheats	☐	☐	☐
Picked on or bullied by other children	☐	☐	☐
Often volunteers to help others (parents, teachers, other children)	☐	☐	☐
Thinks things out before acting	☐	☐	☐
Steals from home, school or elsewhere	☐	☐	☐
Gets on better with adults than with other children	☐	☐	☐
Many fears, easily scared	☐	☐	☐
Sees tasks through to the end, good attention span	☐	☐	☐

Do you have any other comments or concerns?

Appendix 3i: Additional Ethical Considerations

Ethical features	Issue	How resolved	Contingency
3. Recall of personal memories	Programme invites participants to explore their thoughts and feelings in relation to a range of personal and social situations.	Anxiety and negative thoughts are normalised. Programme teaches methods of transforming negative reactions into positive ones.	Participants free not to share worries if they do not wish to. Teacher/EP available to discuss any issues causing concern. 'Posting box'*
8. Procedures likely to change participants' mood or be stressful.	Programme aims to change thoughts and behaviours	Positive self-talk, relaxation and other techniques are taught to reduce stress and anxiety.	Time out procedure in place /adult available for discussion 'Posting box'
10. Tasks to be performed outside school	Weekly home activities	Completion voluntary	
11. Participants under 16	Participants will be Key Stage 2 children	Informed written consent will be obtained from school and parents and the children themselves will be informed about their role in the study and their right to withdraw.	Right to withdraw
12. Participants whose capacity to give consent may be in doubt.	Possible adult/peer pressure to take part	Purpose of study and right to withdraw will be explained to parents, teachers and participants	Right to withdraw
14. Participants recruited from special sources	Participants recruited from mainstream primary schools	As above	Right to withdraw
20. Possible disclosure of confidential information	Participants may reveal confidential information related to a difficult situation	Encourage participants/staff to think about the kinds of problems to share in class.	Teacher or EP available to discuss sensitive cases. Researcher to seek advice through supervision if necessary. Safeguarding procedures.
22. Procedures which might be harmful or distressing to people in a vulnerable state.	Focusing on anxiety-provoking situations may cause distress	Staff to encourage participants to choose manageable problems to work with.	Time out/withdrawal/ adult discussion 'Posting box' Follow up of individual cases if anxiety data gives cause for concern.
23. Procedures from which participants	Possible adult/peer pressure to take part		Right to withdraw data emphasised at

may not feel free to withdraw at any point or may regret taking part in			each data collection point. 'Posting box'
25. Information-gathering on sensitive issues	Self-reports of anxiety and perceptions as a learner; teacher reports of pupil strengths and difficulties	Data to be kept secure and confidential throughout study; anonymity in final report.	Individual cases followed up with teachers and parents/carers as necessary if self-report scores give cause for concern. Referrals made to relevant service providers with parental consent if necessary.
27. Discussion or investigation of personal topics or any procedure in which participants may have an emotional investment.	Consideration of perceptions of self-efficacy and ability to cope with anxiety-provoking situations	Programme teaches positive self-talk and coping strategies which should enhance feelings of self-efficacy	Adult discussion 'Posting box' Follow up of individual cases if anxiety data gives cause for concern.
28. Multiple sessions with the same participant	Programme runs for ten consecutive weeks with two booster sessions.		
29. Lack of back-up/counselling/follow-up arrangements in cases where participants are distressed or embarrassed.	Participants experience a lack of support following discussions of difficult issues		Participants regularly reminded of access to adult 'surgeries' and 'posting box.'

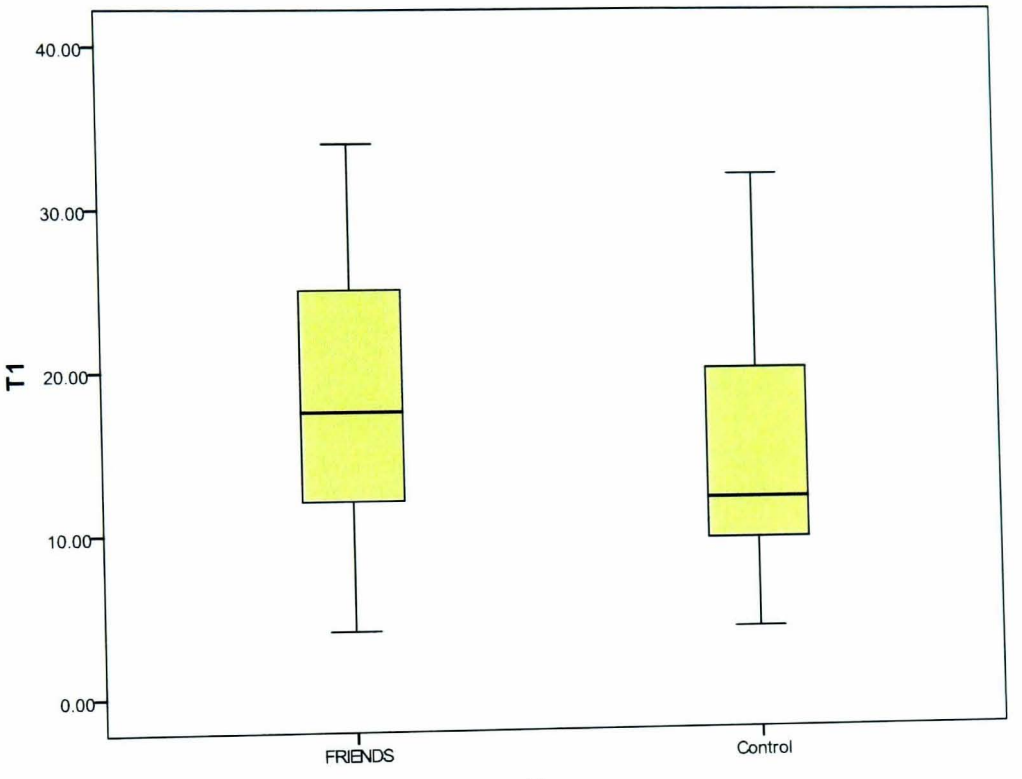
* Post box available for either named or anonymous comments which can be followed up individually or as a whole group as requested.

Appendix 3j: Key observations from the pilot study that influenced the main study implementation

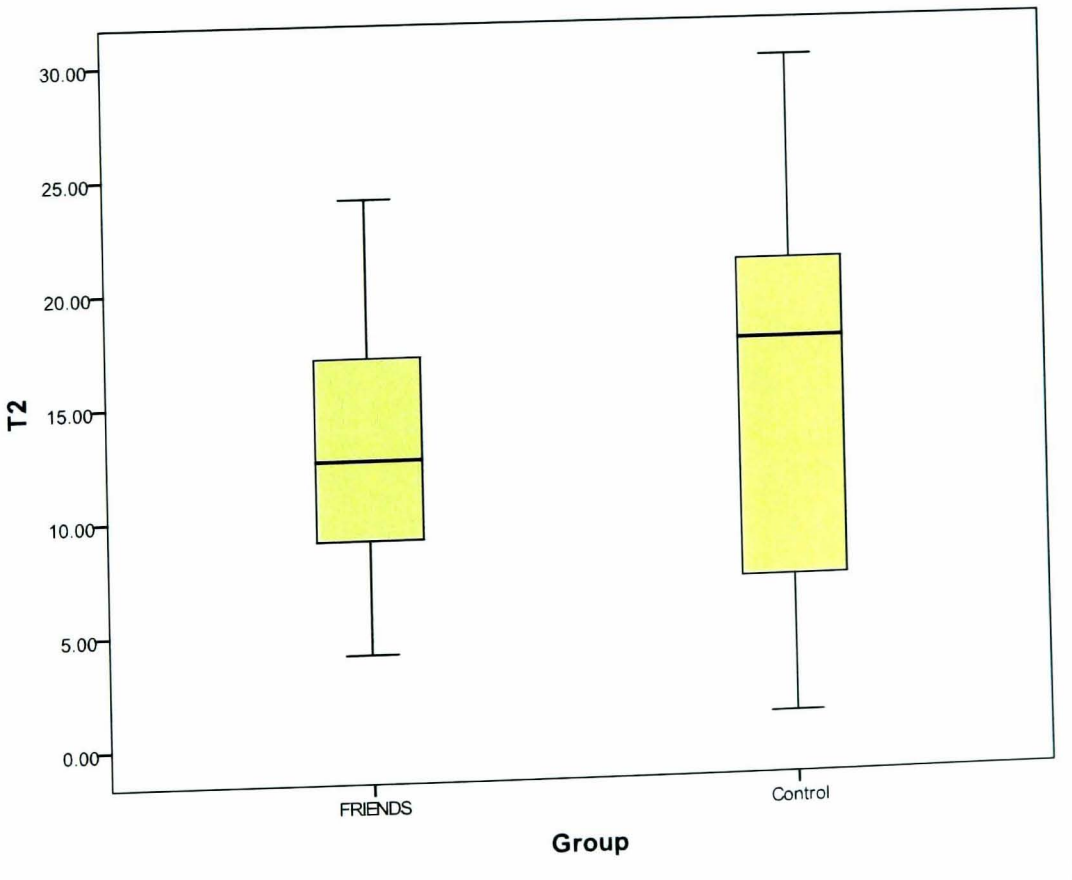
1. The difficulties of accessing parent consent and the benefit of recruiting the family support worker with whom many had an established relationship.
2. The need to involve adults to support poorer readers with the measures and to standardise this process (in the pilot, the scales were administered by different adults to small groups).
3. Reported benefits of having several adults to support during the sessions to enable more individual questioning and clarification about the concepts. In this way, individual children experiencing problems (for example, language, comprehension, social, emotional or attention difficulties) could be targeted for individual support and praise.
4. The teacher noted that children with English as an additional language struggled particularly with the emotional vocabulary and this required extra teaching and reinforcement. Some activities had to be adapted to include more visual cues.
5. Reported benefits of having special stickers and rewards for FRIENDS to reinforce the children's participation and learning.
6. Positive response to the 'worry box' and 'feelings ladder' and their helpfulness in monitoring individual children's concerns.
7. The need to adapt the programme to the children's culture and level of understanding.
8. Reported benefits of having the scrapbook, which could be used more flexibly than the published workbook.
9. Advantages of having the homework club and reinforcing the strategies throughout the week.
10. Descriptive trends between changes on the PI-ED and MALS indicating an association between emotional distress and academic self-perceptions.

Appendix 4a: Box and Whisker Plots to illustrate the distribution and spread of data for both groups at pre and post-test for Emotional Distress.

Emotional Distress Pre-test

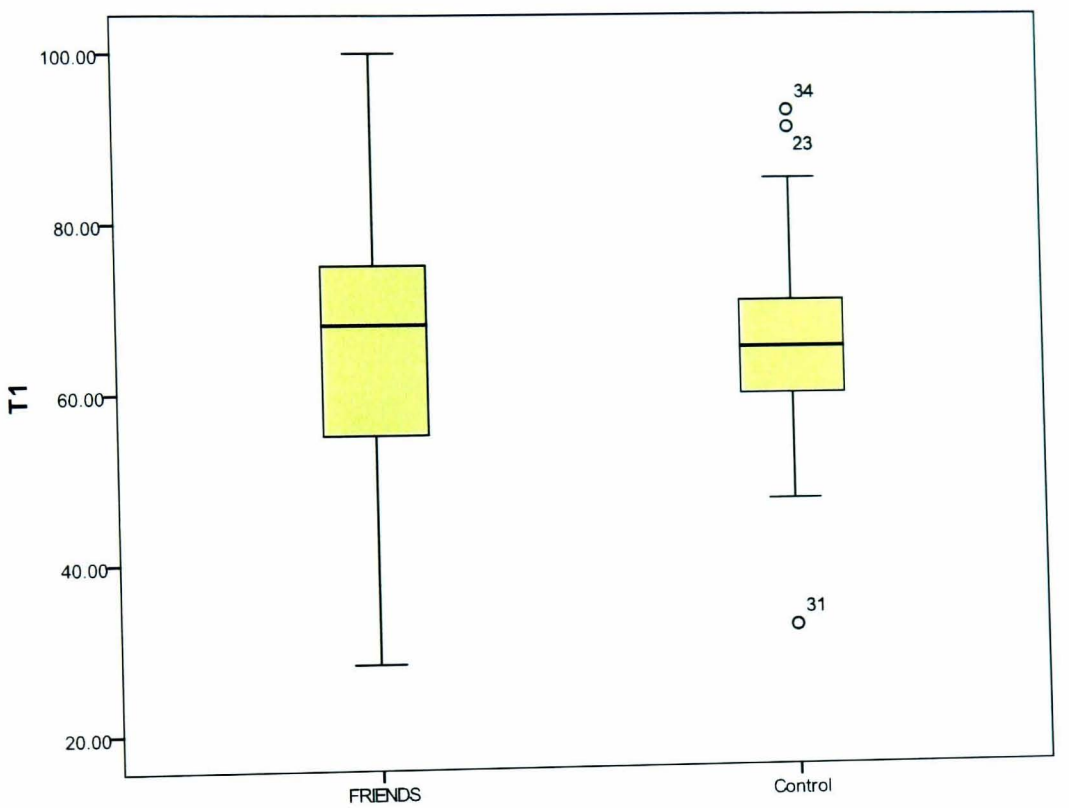


Emotional Distress Post-test

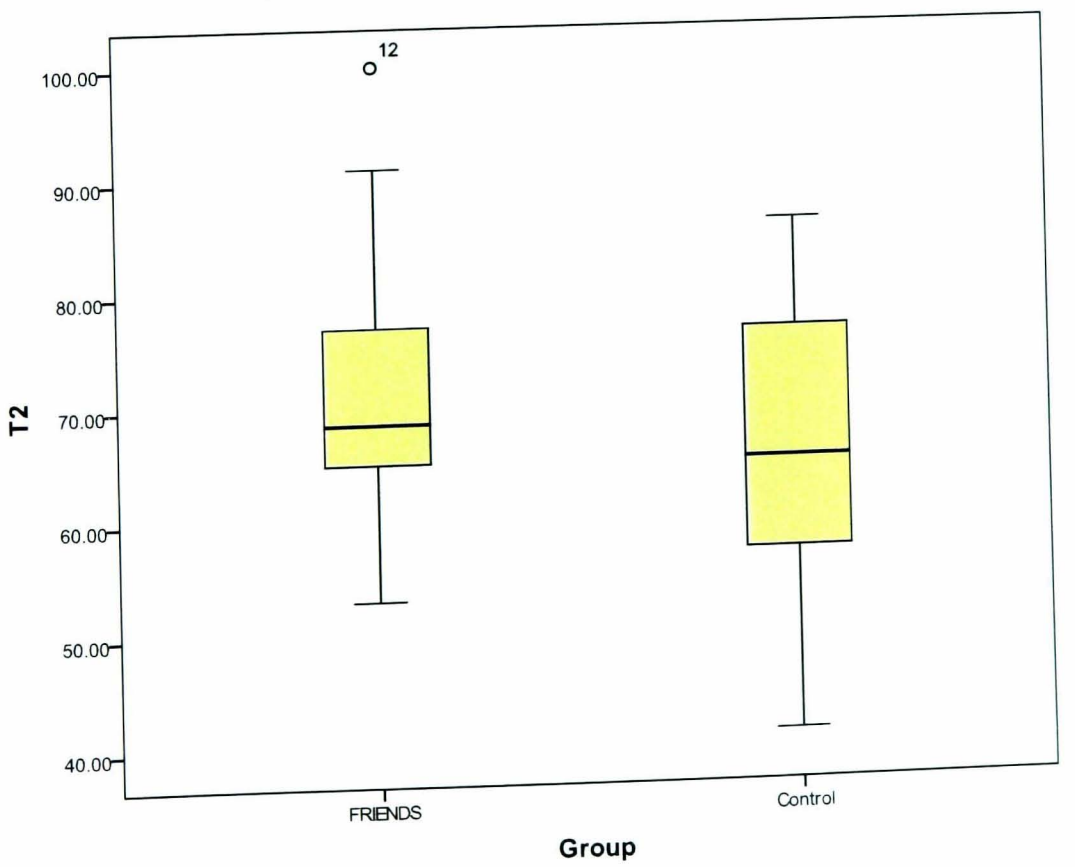


Appendix 4b: Box and Whisker Plots to illustrate the distribution and spread of data for both groups at pre and post-test for Academic Self-perceptions.

Academic Self-Perceptions Pre-test

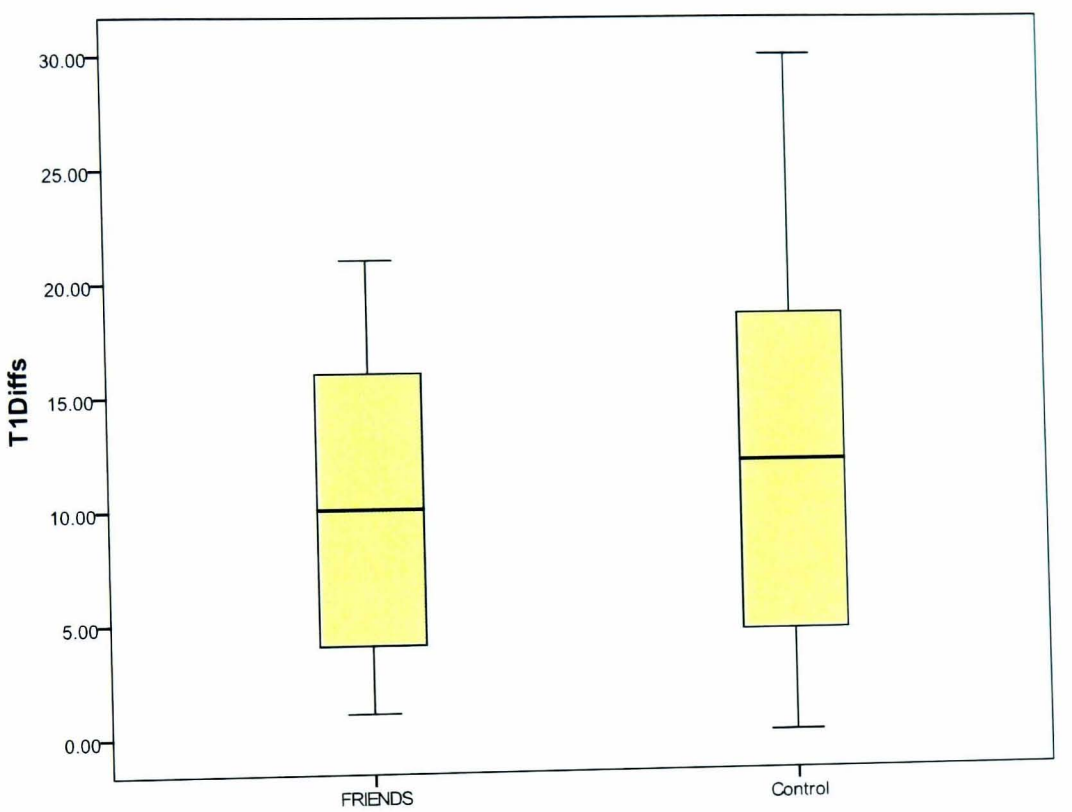


Academic Self-Perceptions Post-test

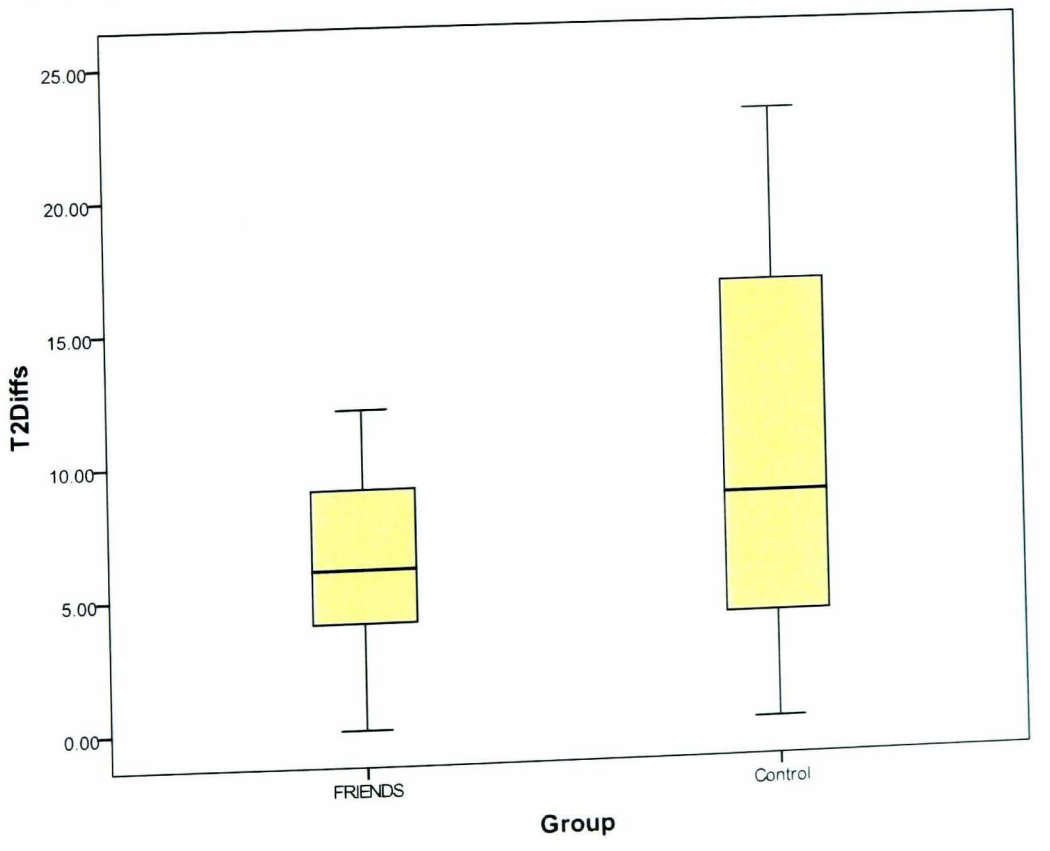


Appendix 4c: Box and Whisker Plots to illustrate the distribution and spread of data for both groups at pre and post-test for Total Difficulties.

Total Difficulties Pre-test



Total Difficulties Post-test



Appendix 4d: Examples of non-normal distribution in the T1 and T2 data necessitating non-parametric analysis

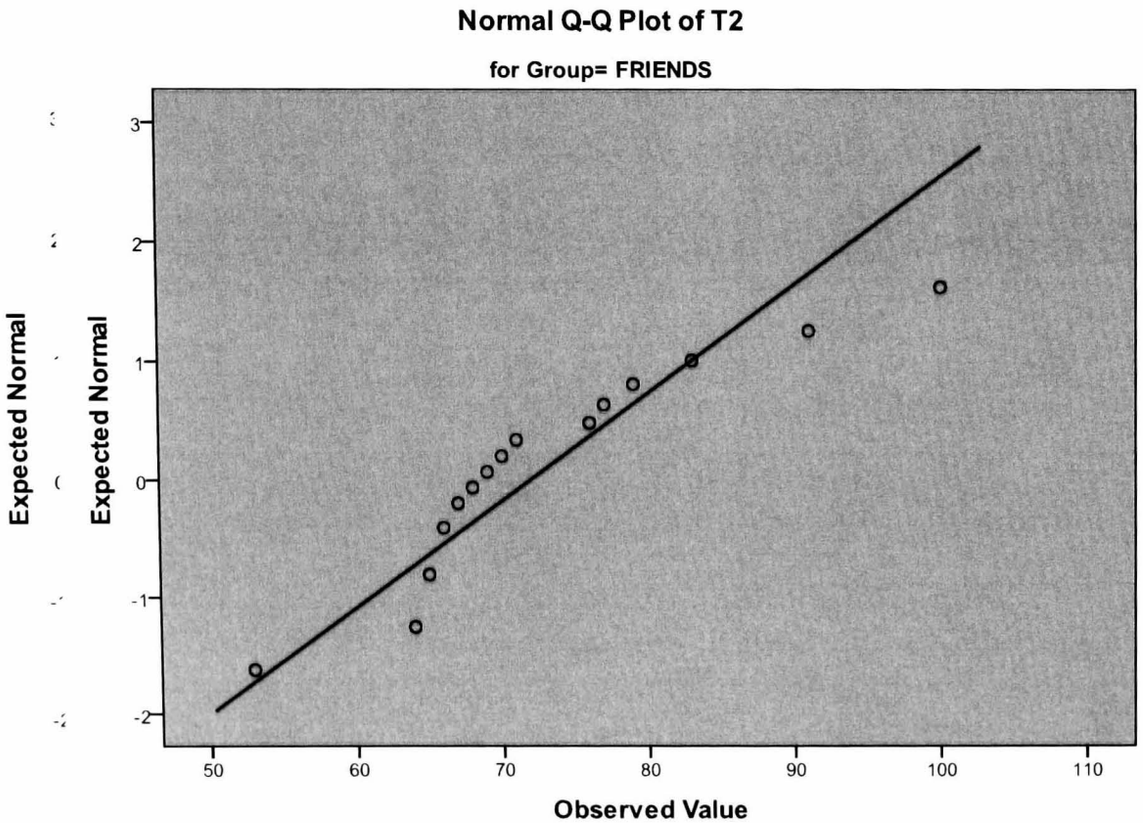
PI-ED Pre-test

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIENDS	.166	.536	.309	-.757	1.038	.729	.965	18	.705
Control	.861	.512	1.68	-.350	.992	.352	.896	20	.034

MALS Post-test

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIENDS	1.093	.536	2.039	1.625	1.038	1.566	.894	18	.045
Control	-.145	.512	.283	-.918	.992	.925	.949	20	.355

Normal Q-Q Plots to illustrate the non-normal distribution of data for the Control Group (PI-ED, T1) and the Intervention Group (MALS, T2).



Appendix 4e: Non parametric analyses of the distribution of scores for both groups at pre-test.

Dependent variable	Group	M (Median)	*U*	*p*	Result	Effect size (r)
Emotional Distress (PI-ED)	IG	17.50	134.0	.178	Not sig.	.22
	CG	12.00				
Academic Self-Perceptions (MALS)	IG	68.00	161.5	.588	Not sig.	.09
	CG	65.00				
Total Difficulties (SDQ)	IG	10.00	194.5	.671	Not sig.	.07
	CG	12.00				
Emotional Symptoms (SDQ)	IG	2.00	186.5	.848	Not sig.	.03
	CG	3.00				
Conduct (SDQ)	IG	.50	223.5	.184	Not sig.	.22
	CG	2.00				
Hyperactivity (SDQ)	IG	4.00	195.5	.649	Not sig.	.07
	CG	6.00				
Peer Problems (SDQ)	IG	1.00	189.0	.787	Not sig.	.04
	CG	2.00				
Prosocial (SDQ)	IG	7.00	175.5	.894	Not sig.	.02
	CG	5.50				

Appendix 4f: Tests of normality and homogeneity of variance for the Emotional Distress change scores

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIENDS	-.73	.54	1.35	1.02	1.04	0.98	.960	18	.601
Control	.26	.51	0.51	-.44	.99	-.44	.971	20	.765

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Gain	FRIENDS	18	-4.6667	8.08775	1.90630
	Control	20	.7500	6.88916	1.54046

		Levene's Test for Equality of Variances	
		F	Sig.
Gain	Equal variances assumed	.439	.512
	Equal variances not assumed		

Conclusion: Normal distribution and homogeneity of variance demonstrated.

Appendix 4g: Tests of normality and homogeneity of variance for the Academic Self-perceptions change scores

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIENDS	.59	.54	1.09	.58	1.04	.56	.937	18	.258
Control	-.51	.51	1.00	.25	.99	.25	.970	20	.748

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Gain	FRIENDS	18	5.8333	13.91254	3.27922
	Control	20	-.1500	10.29192	2.30134

		Levene's Test for Equality of Variances	
		F	Sig.
Gain	Equal variances assumed	.512	.479
	Equal variances not assumed		

Conclusion: Normal distribution and homogeneity of variance demonstrated.

Appendix 4h: Tests of normality and homogeneity of variance for the Emotional Symptoms change scores

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIENDS	1.28	.54	2.37	1.60	1.04	1.54	.846	18	.007
Control	-.66	.51	1.29	-.10	.99	.10	.948	20	.336

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
GainEmotional	FRIENDS	18	-1.5556	1.33823	.31542
	Control	20	-1.1500	2.60111	.58163

		Levene's Test for Equality of Variances	
		F	Sig.
GainEmotional	Equal variances assumed	7.197	.011
	Equal variances not assumed		

Conclusion: Normal distribution and homogeneity of variance NOT demonstrated.

Appendix 4i: Tests of normality and homogeneity of variance for the Conduct change scores

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIEND	-1.77	.54	3.27	1.73	1.04	1.66	.583	18	.000
Control	-1.15	.51	2.25	-.149	.992	.15	.715	20	.000

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
GainConduct	FRIENDS	18	-.5556	1.61690	.38111
	Control	20	-1.0000	1.45095	.32444

		Levene's Test for Equality of Variances	
		F	Sig.
GainConduct	Equal variances assumed	.029	.866
	Equal variances not assumed		

Conclusion: Normal distribution NOT demonstrated.

Appendix 4j: Tests of normality and homogeneity of variance for the Hyperactivity change scores

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIENDS	-.57	.54	1.06	-.62	1.04	0.60	.936	18	.245
Control	.18	.51	0.35	-.26	.99	.26	.964	20	.618

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
GainHyper	FRIENDS	18	-1.8889	2.63213	.62040
	Control	20	-.1500	1.53125	.34240

		Levene's Test for Equality of Variances	
		F	Sig.
GainHyper	Equal variances assumed	5.450	.025
	Equal variances not assumed		

Conclusion: Homogeneity of variance NOT demonstrated.

Appendix 4k: Tests of normality and homogeneity of variance for the Peer Problems change scores

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIENDS	-.15	.54	.28	-.260	1.04	.25	.925	18	.158
Control	-.04	.51	.08	.25	.99	.25	.863	20	.009

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
GainPeer	FRIENDS	18	-.6667	1.28338	.30250
	Control	20	-.1000	1.11921	.25026

		Levene's Test for Equality of Variances	
		F	Sig.
GainPeer	Equal variances assumed	.945	.337
	Equal variances not assumed		

Conclusion: Normal distribution NOT demonstrated.

Appendix 4l: Tests of normality and homogeneity of variance for the Prosocial change scores

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIENDS	1.41	.54	2.61	2.47	1.038	2.38	.876	18	.022
Control	.43	.51	.84	.-79	.99	.80	.922	20	.108

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
GainProsoc	FRIENDS	18	1.7222	2.24409	.52894
	Control	20	1.0000	1.48678	.33245

		Levene's Test for Equality of Variances	
		F	Sig.
GainProsoc	Equal variances assumed	1.471	.233
	Equal variances not assumed		

Conclusion: Normal distribution NOT demonstrated.

Appendix 4m: Tests of normality and homogeneity of variance for the Total Difficulties change scores

Group	Skewness	St. error	Z score (skewness)	Kurtosis	St. error	Z score (kurtosis)	Shapiro-Wilk		
							Statistic	df	Sig.
FRIENDS	-.76	.54	1.41	.33	1.04	.32	.941	18	.303
Control	.41	.51	.80	-.62	.99	.63	.953	20	.423

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
GainDiffs	FRIENDS	18	-4.6667	4.31141	1.01621
	Control	20	-2.4000	3.96564	.88674

		Levene's Test for Equality of Variances	
		F	Sig.
GainDiffs	Equal variances assumed	.014	.907
	Equal variances not assumed		

Conclusion: Normal distribution and homogeneity of variance demonstrated.

Appendix 4n: PASW output for the independent t-test on Emotional Distress change scores

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Gain	FRIENDS	18	-4.6667	8.08775	1.90630
	Control	20	.7500	6.88916	1.54046

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Gain	Equal variances assumed	.439	.512	-2.229	36	.032	-5.41667	2.42992	-10.34477	-.48856
	Equal variances not assumed			-2.210	33.623	.034	-5.41667	2.45092	-10.39959	-.43374

Appendix 4o: PASW output for the independent t-test on Academic Self-Perceptions change scores

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
Gain	FRIENDS	18	5.8333	13.91254	3.27922
	Control	20	-.1500	10.29192	2.30134

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Gain	Equal variances assumed	.512	.479	1.517	36	.138	5.98333	3.94323	-2.01391	13.98058
	Equal variances not assumed			1.494	31.116	.145	5.98333	4.00618	-2.18608	14.15275

Appendix 4p : PASW output for independent t-test on Total Difficulties change scores

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
GainDiffs	FRIENDS	18	-4.6667	4.31141	1.01621
	Control	20	-2.4000	3.96564	.88674

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
GainDiffs	Equal variances assumed	.014	.907	-1.688	36	.100	-2.26667	1.34263	-4.98965	.45631
	Equal variances not assumed			-1.681	34.729	.102	-2.26667	1.34870	-5.00544	.47211

Appendix 4q: PASW output for the independent t-test on Hyperactivity change scores

Group Statistics

	Group	N	Mean	Std. Deviation	Std. Error Mean
GainHyper	FRIENDS	18	-1.8889	2.63213	.62040
	Control	20	-.1500	1.53125	.34240

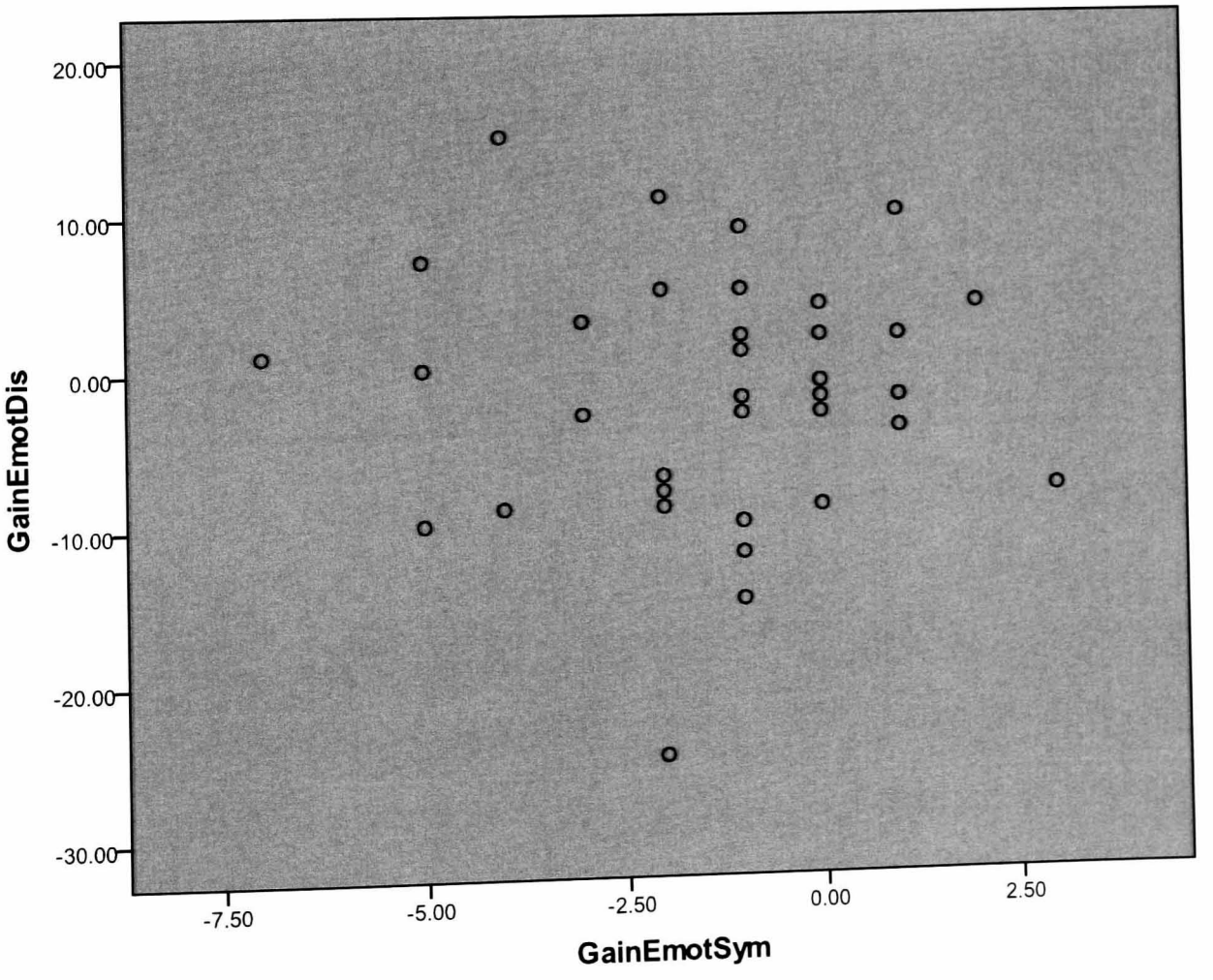
Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
GainHyper	Equal variances assumed	5.450	.025	-2.520	36	.016	-1.73889	.68990	-3.13807	-.33971
	Equal variances not assumed			-2.454	26.716	.021	-1.73889	.70861	-3.19357	-.28421

Appendix 4r: Correlational analysis of the Emotional Distress and Emotional Symptoms change scores.

Correlations

		GainEmotDis	GainEmotSym
GainEmotDis	Pearson Correlation	1	-.021
	Sig. (2-tailed)		.901
	N	38	38
GainEmotSym	Pearson Correlation	-.021	1
	Sig. (2-tailed)	.901	
	N	38	38

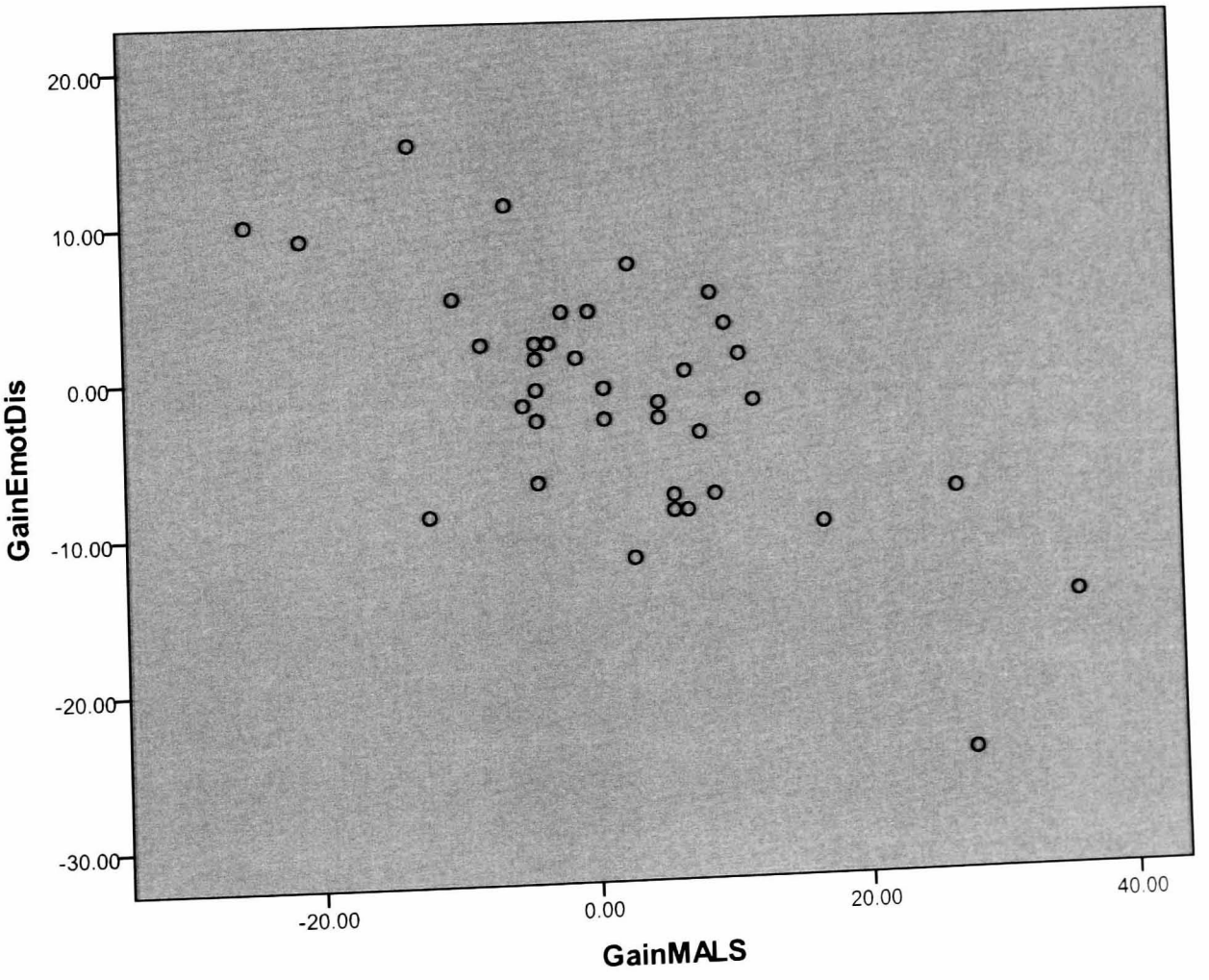


Appendix 4s: Correlational analysis of the Emotional Distress and Academic Self-perceptions change scores

Correlations

		GainEmotDis	GainMALS
GainEmotDis	Pearson Correlation	1	-.693**
	Sig. (2-tailed)		.000
	N	38	38
GainMALS	Pearson Correlation	-.693**	1
	Sig. (2-tailed)	.000	
	N	38	38

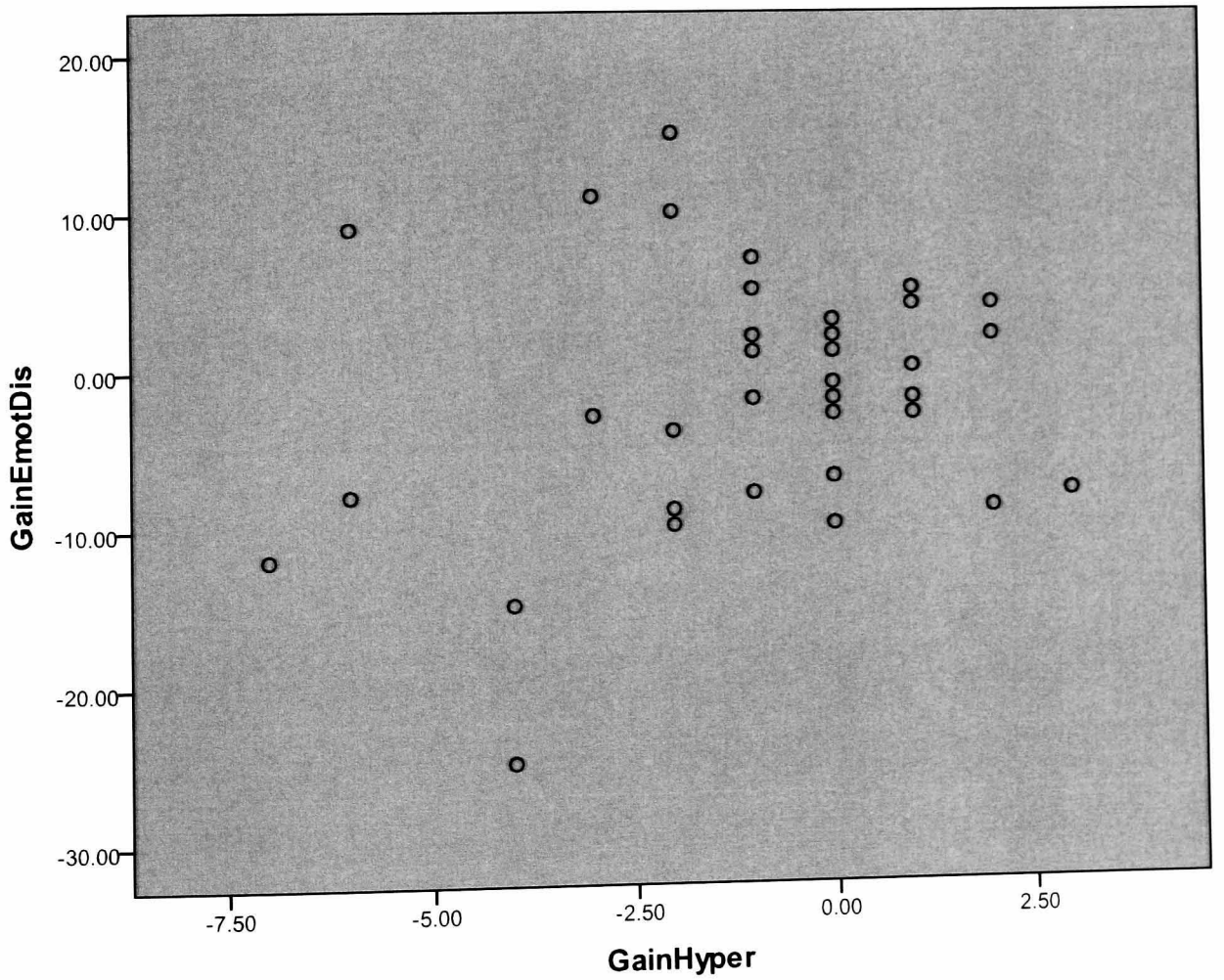
**. Correlation is significant at the 0.01 level (2-tailed).



Appendix 4t: Correlational analysis of the Emotional Distress and Hyperactivity change scores

Correlations

		GainEmotDis	GainHyper
GainEmotDis	Pearson Correlation	1	.196
	Sig. (2-tailed)		.238
	N	38	38
GainHyper	Pearson Correlation	.196	1
	Sig. (2-tailed)	.238	
	N	38	38

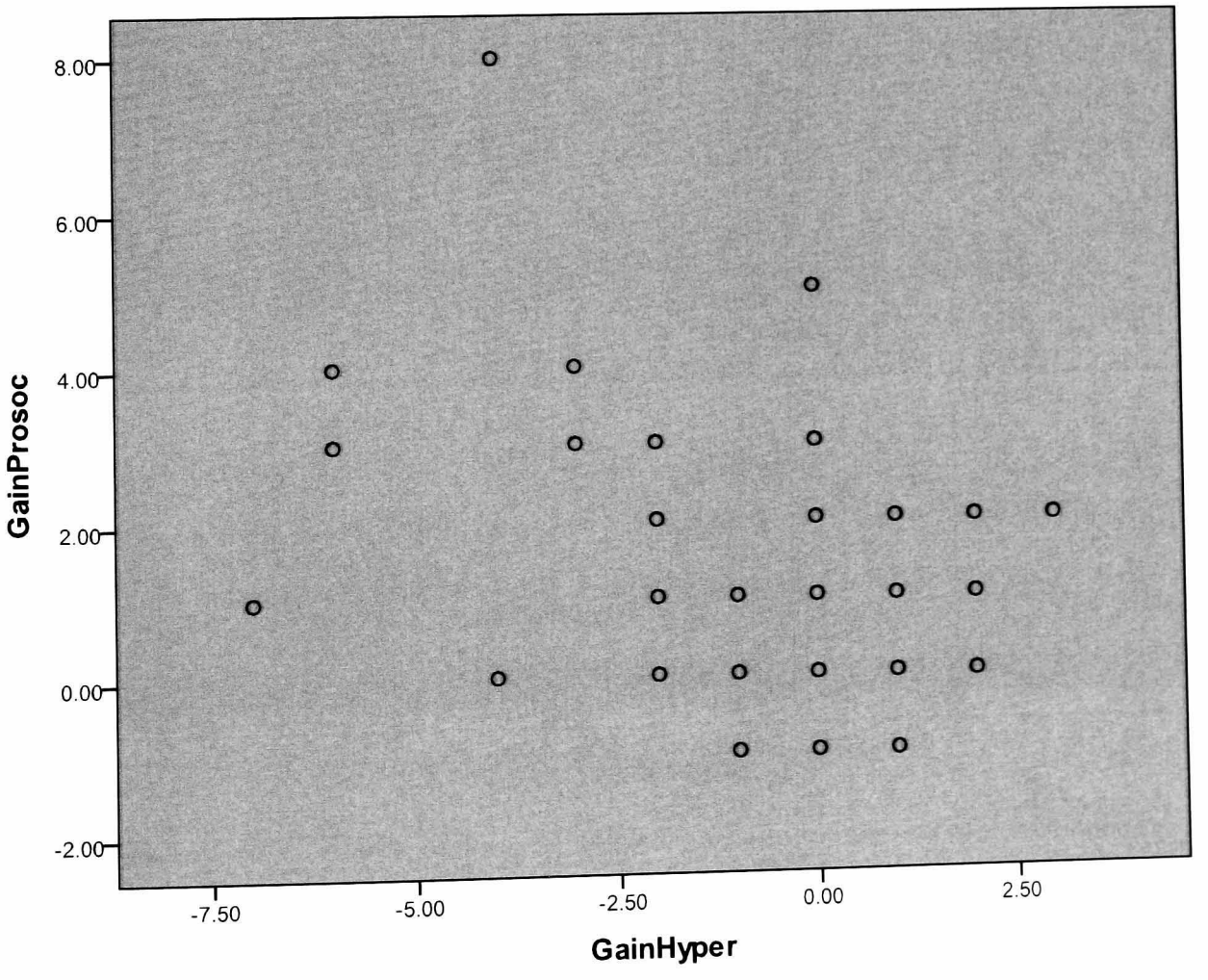


Appendix 4u: Correlational analysis of the Hyperactivity and Prosocial change scores

Correlations

		GainHyper	GainProsoc
GainHyper	Pearson Correlation	1	-.360*
	Sig. (2-tailed)		.026
	N	38	38
GainProsoc	Pearson Correlation	-.360*	1
	Sig. (2-tailed)	.026	
	N	38	38

*. Correlation is significant at the 0.05 level (2-tailed).



Appendix 4v: Correlational analysis of the Emotional Distress and Prosocial change scores

Correlations

		GainProsoc	GainEmotDis
GainProsoc	Pearson Correlation	1	-.021
	Sig. (2-tailed)		.902
	N	38	38
GainEmotDis	Pearson Correlation	-.021	1
	Sig. (2-tailed)	.902	
	N	38	38